

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

rotein - protein search, using sw model

On: August 11, 2004, 13:00:42 ; Search time 45.7143 Seconds  
(without alignments)  
61.807 Million cell updates/sec

US-10-661-207-1

ject score: 47

ience: 1 GPCXXXFIRY 10

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 1586107 seqs, 282547505 residues

l number of hits satisfying chosen parameters: 1586107

um DB seq length: 0

um DB seq length: 2000000000

rocessing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

base : A\_Geneseq\_29Jan04:\*  
1: \_geneseqp1980s:\*  
2: \_geneseqp1990s:\*  
3: \_geneseqp2000s:\*  
4: \_geneseqp2001s:\*  
5: \_geneseqp2002s:\*  
6: \_geneseqp2003as:\*  
7: \_geneseqp2003bs:\*  
8: \_geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

lt	Score	Match	Length	DB	ID	Description
1	42	89.4	58	2	AAR78567	Aar78567 BPTI Kuni
2	42	89.4	58	2	AAR78568	Aar78568 BPTI Kuni
3	42	89.4	58	2	AAR78569	Aar78569 BPTI Kuni
4	41	87.2	10	6	ABU09383	Abu09383 Consensus
5	40	85.1	58	2	AAR99175	Aar99175 Genetical
6	40	85.1	58	2	AAR99176	Aar99176 Genetical
7	37	78.7	57	2	AAW47436	Aaw47436 Aprotinin
8	37	78.7	57	2	AAW47434	Aaw47434 Aprotinin
9	37	78.7	57	2	AAW47432	Aaw47432 Aprotinin
10	37	78.7	57	2	AAW47435	Aaw47435 Aprotinin
11	37	78.7	58	2	AAR78552	Aar78552 Human TFP
12	37	78.7	58	2	AAR99184	Aar99184 Genetical
13	37	78.7	60	2	AAAY08613	Aay08613 Inter-alp
14	37	78.7	128	2	AAW82769	Aaw82769 Bovine bi
15	37	78.7	128	2	AAW82766	Aaw82766 Bovine bi
16	37	78.7	128	2	AAW82768	Aaw82768 Bovine bi
17	37	78.7	128	2	AAW82764	Aaw82764 Bovine bi
18	37	78.7	128	2	AAW82771	Aaw82771 Bovine bi
19	37	78.7	128	2	AAW82767	Aaw82767 Bovine bi
20	37	78.7	128	2	AAW82772	Aaw82772 Bovine bi
21	37	78.7	129	2	AAW82765	Aaw82765 Bovine bi
22	36	76.6	25	1	AAP91296	Aap91296 Amino aci
23	36	76.6	56	1	AAP93399	Aap93399 Aprotinin
24	36	76.6	56	1	AAP93398	Aap93398 Aprotinin
25	36	76.6	56	2	AAR06714	Aar06714 Synthetic

Aay68060 Kunitz pr  
Aau09230 Human KPI  
Aap50699 Sequence  
Aaw47433 Aprotinin  
Aap50700 Sequence  
Aap50701 Sequence  
Aap50698 Sequence  
Aap70336 Sequence  
Aap93401 Aprotinin  
Aap93400 Aprotinin  
Aap92306 Aprotinin  
Aap93329 Sequence  
Aap93331 Sequence  
Aap93330 Sequence  
Aar27373 Bovine pa  
Aar39801 Kunitz-ty  
Aar62524 Pancreati  
Aar47179 BPTI. 3/2  
Aar78426 Bovine pa  
Aar81885 Mature bo

ALIGNMENTS

RESULT 1  
AAR78567  
ID AAR78567 standard; peptide; 58 AA.  
XX  
AC AAR78567;  
XX  
DT 01-MAR-1996 (first entry)  
XX  
DE BPTI Kunitz domain derivative DPI-7.1.  
XX

KW Human; lipoprotein-associated coagulation inhibitor; peptide library;  
inhibitor; plasmin; bovine; pancreatic trypsin inhibitor; Kunitz domain;  
KW fibrinolysis; fibrinogenolysis; bleeding; thrombolytic.  
XX

OS Synthetic.

XX  
FH Key Location/Qualifiers  
FT Misc-difference 18  
FT /note= "residue change: Ile to Phe"

XX WO9518830-A2.

XX 13-JUL-1995.

XX 11-JAN-1995; 95WO-US000298.

PR 11-JAN-1994; 94US-00179658.

PR 10-MAR-1994; 94US-00208265.

XX (PROT-) PROTEIN ENG CORP.

XX Markland W, Ladner RC;

XX WPI; 1995-255042/33.

PT Novel plasmin inhibiting protein comprising a Kunitz Domain - useful to  
prevent/treat disorders attributable to excess plasmin activity.

XX Example 2; Page 36; 59pp; English.

CC The peptides AAR78435-R78570 are derivatives of the Kunitz domains from a  
variety of plasmin inhibitors e.g. the human lipoprotein-associated  
coagulation inhibitor (LACI) Kunitz domains 1, 2 or 3. The peptides were  
designed based on the Kunitz domains and are named Designed plasmid  
Inhibitor (DPI). This peptide is designed based on the bovine pancreatic  
trypsin inhibitor (BPTI) Kunitz domain (KuDom) (AAR78426). The peptides  
can be used to prevent or treat a clinical condition exacerbated by  
plasmin e.g. inappropriate fibrinolysis or fibrinogenolysis, excessive

CC bleeding associated with thrombolytics

XX Sequence 58 AA;

Query Match 89.4%; Score 42; DB 2; Length 58;  
Best Local Similarity 70.0%; Pred. No. 1.1;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| ||||  
Db 12 GPCXARFIRY 21

RESULT 2

AAR78568  
ID AAR78568 standard; peptide; 58 AA.

XX  
AC AAR78568;

XX  
DT 01-MAR-1996 (first entry)

DE BPTI Kunitz domain derivative DPI-7.2.

XX Human; lipoprotein-associated coagulation inhibitor; peptide library;  
KW inhibitor; plasmin; bovine; pancreatic trypsin inhibitor; Kunitz domain;  
KW fibrinolysis; fibrinogenolysis; bleeding; thrombolytic.

XX  
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 15 /note= "residue change: Lys to Arg"

FT Misc-difference 18 /note= "residue change: Ile to Phe"

FT Misc-difference 39 /note= "residue change: Arg to Gly"

XX WO9518830-A2.

PN 13-JUL-1995.

XX 11-JAN-1995; 95WO-US000298.

XX 11-JAN-1994; 94US-00179658.

PR 10-MAR-1994; 94US-00208265.

XX (PROT-) PROTEIN ENG CORP.

XX Markland W, Ladner RC;

XX WPI; 1995-255042/33.

XX Novel plasmin inhibiting protein comprising a Kunitz Domain - useful to prevent/treat disorders attributable to excess plasmin activity.

PS Example 2; Page 36; 59pp; English.

XX The peptides AAR78435-R78570 are derivatives of the Kunitz domains from a variety of plasmin inhibitors e.g. the human lipoprotein-associated coagulation inhibitor (LACI) Kunitz domains 1, 2 or 3. The peptides were designed based on the Kunitz domains and are named Designed Plasmid Inhibitor (DPI). This peptide is designed based on the bovine pancreatic trypsin inhibitor (BPTI) Kunitz domain (KuDom) (AAR78426). The peptides can be used to prevent or treat a clinical condition exacerbated by plasmin e.g. inappropriate fibrinolysis or fibrinogenolysis, excessive bleeding associated with thrombolytics

XX Sequence 58 AA;

Query Match

Best Local Similarity 89.4%; Score 42; DB 2; Length 58;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| ||||

Db 12 GPCRARFIRY 21

RESULT 3

AAR78569  
ID AAR78569 standard; peptide; 58 AA.

XX AAR78569;

XX 01-MAR-1996 (first entry)

DE BPTI Kunitz domain derivative DPI-7.3.

XX Human; lipoprotein-associated coagulation inhibitor; peptide library;  
KW inhibitor; plasmin; bovine; pancreatic trypsin inhibitor; Kunitz domain;  
KW fibrinolysis; fibrinogenolysis; bleeding; thrombolytic.

XX  
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 15 /note= "residue change: Lys to Arg"

FT Misc-difference 18 /note= "residue change: Ile to Phe"

FT Misc-difference 39 /note= "residue change: Arg to Gly"

XX WO9518830-A2.

PN 13-JUL-1995.

XX 11-JAN-1995; 95WO-US000298.

XX 11-JAN-1994; 94US-00179658.

PR 10-MAR-1994; 94US-00208265.

XX (PROT-) PROTEIN ENG CORP.

XX Markland W, Ladner RC;

XX WPI; 1995-255042/33.

XX Novel plasmin inhibiting protein comprising a Kunitz Domain - useful to prevent/treat disorders attributable to excess plasmin activity.

PS Example 2; Page 36; 59pp; English.

XX The peptides AAR78435-R78570 are derivatives of the Kunitz domains from a variety of plasmin inhibitors e.g. the human lipoprotein-associated coagulation inhibitor (LACI) Kunitz domains 1, 2 or 3. The peptides were designed based on the Kunitz domains and are named Designed Plasmid Inhibitor (DPI). This peptide is designed based on the bovine pancreatic trypsin inhibitor (BPTI) Kunitz domain (KuDom) (AAR78426). The peptides can be used to prevent or treat a clinical condition exacerbated by plasmin e.g. inappropriate fibrinolysis or fibrinogenolysis, excessive bleeding associated with thrombolytics

XX Sequence 58 AA;

Query Match

Best Local Similarity 89.4%; Score 42; DB 2; Length 58;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| ||||  
Db 12 GPCRARFIRY 21

RESULT 4

ABU09383  
ID ABU09383 standard; peptide; 10 AA.

XX  
AC ABU09383;

17-JUL-2003 (first entry)

Consensus peptide sequence for mammalian Kunitz inhibitor.

Protein purification; protein isolation; recombinant fusion polypeptide; enzyme/tagged-peptide binding; fusion protein; tag peptide; recombinant capture protein; mammalian; Kunitz inhibitor.

Mammalia.

Synthetic.

Key Location/Qualifiers

Misc-difference 4

/label= Lys, Arg

Misc-difference 5

/label= Gly, Ala

Misc-difference 6

/label= Arg, Gly, Ser

EPI281716-A2.

05-FEB-2003.

06-JUL-2002; 2002EP-00015128.

10-JUL-2001; 2001US-00901996.

(HOFF ) ROCHE DIAGNOSTICS GMBH.

(HOFF ) HOFFMANN LA ROCHE & CO AG F.

Dwulet FE, Balgobin NG, Mccarthy RC;

WPI; 2003-334999/32.

Purifying or isolating a recombinant fusion peptide, comprises forming a fusion peptide comprising a tag peptide covalently attached to a polypeptide and contacting the fusion peptide with an enzyme or modified enzyme.

Claim 9; Page 6; 17pp; English.

The present invention relates to a method for purifying or isolating a recombinant fusion polypeptide based on enzyme/tagged-peptide binding. The method comprises forming a fusion protein comprising a tag peptide sequence covalently attached to a polypeptide sequence, and contacting the fusion peptide with an enzyme or modified enzyme that specifically binds to the tag peptide sequence to form a complex between the enzyme or modified enzyme and the fusion peptide. The method of the invention is useful for purifying or isolating recombinant fusion polypeptides. The present method of using enzymes as capture agents offers several advantages not currently found in prior art. For example (a) it is possible to select natural or recombinant capture proteins with modified binding sites that will have different affinities for the same tag, (b) it is possible to modify the amino acid sequence of the tag to generate high, medium and low affinity peptide tags for use in different applications with the same capture protein, and (c) depending upon the application, natural or recombinant capture proteins with increased or decreased resistance to denaturation can also be prepared. Such versatility is amenable to the development of a variety of standardised binding and elution conditions for the isolation of tagged proteins or their complexes. ABU0383-ABU09395 represent peptides that may be used as tag peptides in the method of the present invention

Sequence 10 AA;

Query Match 87.2%; Score 41; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GPCXXXFIRY 10

|||||

1 GPCXXXFIRY 10

RESULT 5

AAR99175

ID AAR99175 standard; protein; 58 AA.

XX AAR99175;

AC AAR99175;

XX 12-FEB-1997 (first entry)

XX Genetically engineered aprotinin-like Kunitz domain (DPI.2.2).  
XX Aprotinin; Kunitz domain; human neutrophil elastase; hNE;  
XX connective tissue; alpha 1 protease inhibitor; API; neutrophil;  
XX alaph antitrypsin; respiratory disorder; cystic fibrosis;  
XX smokers emphysema.

OS Synthetic.

XX WO9620278-A2.

XX 04-JUL-1996.

XX 15-DEC-1995; 95WO-US016349.

XX 16-DEC-1994; 94US-00358160.

XX (PROT-) PROTEIN ENG CORP.

XX Ley AC, Ladner RC, Guterman SK, Roberts BL, Markland W, Kent RB;

XX WPI; 1996-321851/32.

XX New engineered inhibitors of human neutrophil elastase - contg. aprotinin  
XX -like Kunitz domain for treating, e.g. cystic fibrosis or other  
XX respiratory disorders.

PS Claim 1; Page 50; 105pp; English.

XX Genetically engineered human derived Kunitz domains can be used to  
XX inhibit human neutrophil elastase, an enzyme involved in the elimination  
XX of pathogens and the restructuring of connective tissue. In cases of  
XX reduction of the circulating alpha-1-protease inhibitor (API or alaph  
XX antitrypsin), or the inactivation of API by oxidation (smokers  
XX emphysema), extensive destruction of the lung tissue may result from  
XX uncontrolled elastolytic activity of human neutrophil elastase. Other  
XX respiratory disorders such as cystic fibrosis are thought to be caused by  
XX human neutrophil elastase release by neutrophils. The genetically  
XX engineered human derived Kunitz domains can be used to treat such  
XX respiratory disorders. See AAR99146-R99211

SQ Sequence 58 AA;

Query Match 85.1%; Score 40; DB 2; Length 58;  
Best Local Similarity 60.0%; Pred. No. 2.7;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GPCXXXFIRY 10

|||||

Db 12 GPCIALFLRY 21

RESULT 6

AAR99176

ID AAR99176 standard; protein; 58 AA.

XX AAR99176;

XX 12-FEB-1997 (first entry)

XX Genetically engineered aprotinin-like Kunitz domain (DPI.2.3).

XX Aprotinin; Kunitz domain; human neutrophil elastase; hNE;

KW connective tissue; alpha 1 protease inhibitor; API; neutrophil;  
KW alpha1 antitrypsin; respiratory disorder; cystic fibrosis;  
KW smokers emphysema.  
XX  
OS Synthetic.  
XX  
PN WO9620278-A2.  
XX  
PD 04-JUL-1996.  
XX  
PF 15-DEC-1995; 95WO-US016349.  
XX  
PR 16-DEC-1994; 94US-00358160.  
XX  
PA (PROT-) PROTEIN ENG CORP.  
XX  
PI Ley AC, Ladner RC, Guterman SK, Roberts BL, Markland W, Kent RB;  
XX  
DR WPI; 1996-321851/32.  
XX  
PT New engineered inhibitors of human neutrophil elastase - contg. aprotinin  
PT -like Kunitz domain for treating, e.g. cystic fibrosis or other  
PT respiratory disorders.  
XX  
PS Claim 1; Page 50; 105pp; English.  
XX  
CC Genetically engineered human derived Kunitz domains can be used to  
CC inhibit human neutrophil elastase, an enzyme involved in the elimination  
CC of pathogens and the restructuring of connective tissue. In cases of  
CC reduction of the circulating alpha-1-protease inhibitor (API or alpha1  
CC antitrypsin), or the inactivation of API by oxidation (smokers  
CC emphysema), extensive destruction of the lung tissue may result from  
CC uncontrolled elastolytic activity of human neutrophil elastase. Other  
CC respiratory disorders such as cystic fibrosis are thought to be caused by  
CC human neutrophil elastase release by neutrophils. The genetically  
CC engineered human derived Kunitz domains can be used to treat such  
CC respiratory disorders. See AAR99146-R99211  
XX  
SQ Sequence 58 AA;

Query Match 85.1%; Score 40; DB 2; Length 58;  
Best Local Similarity 60.0%; Pred. No. 2.7;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
Db 12 GPCIAFFLRY 21

RESULT 7  
AAW47436  
ID AAW47436 standard; peptide; 57 AA.  
XX  
AC AAW47436;  
XX  
DT 26-JUN-1998 (first entry)  
XX  
DE Aprotinin variant 5.  
XX  
KW Aprotinin variant; serine protease inhibitor.  
XX

OS Homo sapiens.  
OS Synthetic.  
XX  
PN EP821007-A2.  
XX  
PD 28-JAN-1998.  
XX  
PF 14-JUL-1997; 97EP-00111980.  
XX  
PR 25-JUL-1996; 96DE-01029982.  
XX  
PA (FARB ) BAYER AG.

XX Schroeder W, Bjorn S, Norris K, Diness V, Norkov-Lauritsen L;  
PI Christensen ND;  
XX  
DR WPI; 1998-102619/10.  
XX  
PT Aprotinin variants prepared by recombinant DNA techniques - useful as  
PT protease inhibitors in surgery and medicine.  
XX  
PS Claim 5; Fig 1; 19pp; German.  
XX  
CC The present sequence is an aprotinin variant, which can be used as a  
CC serine protease inhibitor, e.g. in surgery to reduce blood loss, to treat  
CC multiple trauma, shock, pain, oedema, stroke and inflammation and to  
CC inhibit invasive tumour growth and metastasis, inflammation and to  
CC and risk of haemorrhage in dialysis therapy and artificial organs.  
CC Compared with wild type aprotinin, the variant is less reactive with anti  
CC -aprotinin antibodies, less immunogenic, causes less histamine release  
CC from blood cells, exhibits reduced renal accumulation and has improved Ki  
CC values  
XX  
SQ Sequence 57 AA;

Query Match 78.7%; Score 37; DB 2; Length 57;  
Best Local Similarity 60.0%; Pred. No. 9.8;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
Db 11 GPCRASIIRY 20

RESULT 8  
AAW47434  
ID AAW47434 standard; peptide; 57 AA.  
XX  
AC AAW47434;  
XX  
DT 26-JUN-1998 (first entry)  
XX  
DE Aprotinin variant 3.  
XX  
KW Aprotinin variant; serine protease inhibitor.  
XX

OS Homo sapiens.  
OS Synthetic.  
XX  
PN EP821007-A2;  
XX  
PD 28-JAN-1998.  
XX  
PF 14-JUL-1997; 97EP-00111980.  
XX  
PR 25-JUL-1996; 96DE-01029982.  
XX  
PA (FARB ) BAYER AG.  
XX

PI Schroeder W, Bjorn S, Norris K, Diness V, Norkov-Lauritsen L;  
PI Christensen ND;  
XX  
DR WPI; 1998-102619/10.  
XX  
PT Aprotinin variants prepared by recombinant DNA techniques - useful as  
PT protease inhibitors in surgery and medicine.  
XX  
PS Claim 5; Fig 1; 19pp; German.  
XX

CC The present sequence is an aprotinin variant, which can be used as a  
CC serine protease inhibitor, e.g. in surgery to reduce blood loss, to treat  
CC multiple trauma, shock, pain, oedema, stroke and inflammation and to  
CC inhibit invasive tumour growth and metastasis, inflammation, coagulation  
CC and risk of haemorrhage in dialysis therapy and artificial organs.  
CC Compared with wild type aprotinin, the variant is less reactive with anti



-aprotinin antibodies, less immunogenic, causes less histamine release from blood cells, exhibits reduced renal accumulation and has improved Ki values

Sequence 57 AA;

Query Match 78.7%; Score 37; DB 2; Length 57;  
Best Local Similarity 60.0%; Pred. No. 9.8;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GPCXXXFIRY 10  
||| |||  
11 GPCRAAIIRY 20

JLT 9  
17432

AAW47432 standard; peptide; 57 AA.

AAW47432;

26-JUN-1998 (first entry)

Aprotinin variant 1.

Aprotinin variant; serine protease inhibitor.

Homo sapiens.  
Synthetic.

EP821007-A2.

28-JAN-1998.

14-JUL-1997; 97EP-00111980.

25-JUL-1996; 96DE-01029982.

(FARB ) BAYER AG.

Schroeder W, Bjorn S, Norris K, Diness V, Norkov-Lauritsen L;  
Christensen ND;

WPI; 1998-102619/10.

Aprotinin variants prepared by recombinant DNA techniques - useful as  
protease inhibitors in surgery and medicine.

Claim 5; Page 13; 19pp; German.

The present sequence is an aprotinin variant, which can be used as a  
serine protease inhibitor, e.g. in surgery to reduce blood loss, to treat  
multiple trauma, shock, pain, oedema, stroke and inflammation and to  
inhibit invasive tumour growth and metastasis, inflammation, coagulation  
and risk of haemorrhage in dialysis therapy and artificial organs.  
Compared with wild type aprotinin, the variant is less reactive with anti  
-aprotinin antibodies, less immunogenic, causes less histamine release  
from blood cells, exhibits reduced renal accumulation and has improved Ki  
values

Sequence 57 AA;

Query Match 78.7%; Score 37; DB 2; Length 57;  
Best Local Similarity 60.0%; Pred. No. 9.8;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GPCXXXFIRY 10  
||| |||  
11 GPCRAAIIRY 20

JLT 10  
17435

ID AAW47435 standard; peptide; 57 AA.  
XX AAW47435;  
AC  
XX  
DT 26-JUN-1998 (first entry)  
XX  
DE Aprotinin variant 4.  
XX  
KW Aprotinin variant; serine protease inhibitor.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN EP821007-A2.  
XX  
PD 28-JAN-1998.  
XX  
PF 14-JUL-1997; 97EP-00111980.  
XX  
PR 25-JUL-1996; 96DE-01029982.  
XX  
PA (FARB ) BAYER AG.

PI Schroeder W, Bjorn S, Norris K, Diness V, Norkov-Lauritsen L;  
PI Christensen ND;

XX WPI; 1998-102619/10.

PT Aprotinin variants prepared by recombinant DNA techniques - useful as  
protease inhibitors in surgery and medicine.

PS Claim 5; Fig 1; 19pp; German.

XX The present sequence is an aprotinin variant, which can be used as a  
serine protease inhibitor, e.g. in surgery to reduce blood loss, to treat  
multiple trauma, shock, pain, oedema, stroke and inflammation and to  
inhibit invasive tumour growth and metastasis, inflammation, coagulation  
and risk of haemorrhage in dialysis therapy and artificial organs.  
CC Compared with wild type aprotinin, the variant is less reactive with anti  
-aprotinin antibodies, less immunogenic, causes less histamine release  
from blood cells, exhibits reduced renal accumulation and has improved Ki  
values

XX Sequence 57 AA;

Query Match 78.7%; Score 37; DB 2; Length 57;  
Best Local Similarity 60.0%; Pred. No. 9.8;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| |||  
Db 11 GPCRAAIIRY 20

RESULT 11

AAR78552

ID AAR78552 standard; peptide; 58 AA.

XX AAR78552;

XX 01-MAR-1996 (first entry)

XX Human TFPI-2 domain 3 KUDOM derivative DPI-3.3.1.

XX Human; lipoprotein-associated coagulation inhibitor; peptide library;  
KW inhibitor; plasmin; bovine; pancreatic trypsin inhibitor; Kunitz domain;  
KW fibrinolysis; fibrinogenolysis; bleeding; thrombolytic.

XX Synthetic.

OS Key Location/Qualifiers

FT Misc-difference 11

FT /note= "residue change: Glu to Thr"

FT Misc-difference 13 /note= "residue change: Leu to Pro"  
FT Misc-difference 15 /note= "residue change: Ser to Arg"  
FT Misc-difference 17 /note= "residue change: Asn to Arg"  
FT Misc-difference 18 /note= "residue change: Val to Phe"  
FT Misc-difference 36 /note= "residue change: Thr to Gly"  
XX  
XX  
PN WO9518830-A2.  
XX  
PD 13-JUL-1995.  
XX  
PF 11-JAN-1995; 95WO-US000298.  
XX  
PR 11-JAN-1994; 94US-00179658.  
PR 10-MAR-1994; 94US-00208265.  
XX  
PA (PROT-) PROTEIN ENG CORP.  
XX  
PI Markland W, Ladner RC;  
XX  
DR WPI; 1995-255042/33.  
XX  
PT Novel plasmin inhibiting protein comprising a Kunitz Domain - useful to  
PT prevent/treat disorders attributable to excess plasmin activity.  
XX  
XX  
PS Claim 3; Page 34; 59pp; English.  
XX  
CC The peptides AAR78435-R78570 are derivatives of the Kunitz domains from a  
CC variety of plasmin inhibitors e.g. the human lipoprotein-associated  
CC coagulation inhibitor (LACI) Kunitz domains 1, 2 or 3. The peptides were  
CC designed based on the Kunitz domains and are named Designed Plasmid  
CC Inhibitor (DPI). This peptide is designed based on the human TFPI-2  
CC domain 3 Kunitz domain (KuDom) (AAR78551). The peptides can be used to  
CC prevent or treat a clinical condition exacerbated by plasmin e.g.  
CC inappropriate fibrinolysis or fibrinogenolysis, excessive bleeding  
CC associated with thrombolytics  
XX  
SQ Sequence 58 AA;

Query Match 78.7%; Score 37; DB 2; Length 58;  
Best Local Similarity 60.0%; Pred. No. 9.9;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| |  
Db 12 GPCRRFFTRY 21

RESULT 12  
AAR99184  
ID AAR99184 standard; protein; 58 AA.  
XX  
AC AAR99184;  
XX  
DT 12-FEB-1997 (first entry)  
XX  
DE Genetically engineered aprotinin-like Kunitz domain (DPI.4.3).

XX Aprotinin; Kunitz domain; human neutrophil elastase; hNE;  
KW connective tissue; alpha 1 protease inhibitor; API; neutrophil;  
KW alpha1 antitrypsin; respiratory disorder; cystic fibrosis;  
KW smokers emphysema.  
XX  
OS Synthetic.  
XX  
PN WO9620278-A2.  
XX  
PD 04-JUL-1996.  
XX

PF 15-DEC-1995; 95WO-US016349.  
XX  
PR 16-DEC-1994; 94US-00358160.  
XX  
PA (PROT-) PROTEIN ENG CORP.  
XX  
PI Ley AC, Ladner RC, Guterman SK, Roberts BL, Markland W, Kent RB;  
XX  
DR WPI; 1996-321851/32.  
XX  
PT New engineered inhibitors of human neutrophil elastase - contg. aprotinin  
PT -like Kunitz domain for treating, e.g. cystic fibrosis or other  
PT respiratory disorders.  
XX  
PS Claim 1; Page 51; 105pp; English.  
XX  
CC Genetically engineered human derived Kunitz domains can be used to  
CC inhibit human neutrophil elastase, an enzyme involved in the elimination  
CC of pathogens and the restructuring of connective tissue. In cases of  
CC reduction of the circulating alpha-1-protease inhibitor (API or alpha  
CC antitrypsin), or the inactivation of API by oxidation (smokers  
CC emphysema), extensive destruction of the lung tissue may result from  
CC uncontrolled elastolytic activity of human neutrophil elastase. Other  
CC respiratory disorders such as cystic fibrosis are thought to be caused by  
CC human neutrophil elastase release by neutrophils. The genetically  
CC engineered human derived Kunitz domains can be used to treat such  
CC respiratory disorders. See AAR99146-R99211  
XX  
SQ Sequence 58 AA;

Query Match 78.7%; Score 37; DB 2; Length 58;  
Best Local Similarity 60.0%; Pred. No. 9.9;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| |  
Db 12 GPCIAFFTRY 21

RESULT 13  
AAY08613  
ID AAY08613 standard; protein; 60 AA.  
XX  
AC AAY08613;  
XX  
DT 05-AUG-1999 (first entry)  
XX  
DE Inter-alpha-trypsin inhibitor (BPI type) protein fragment TIBOR.  
XX  
KW APP; beta-amyloid precursor protein; human; transgenic mice; pathology;  
KW Alzheimer's disease; model; therapeutic compound; brain; mechanism;  
KW nerve tissue specific promoter; synthesis; inhibitor; deposition;  
KW plaque formation; treatment.  
XX  
OS Unidentified.  
XX  
PN US5912410-A.  
XX  
PD 15-JUN-1999.  
XX  
PF 13-APR-1995; 95US-00422333.  
XX  
PR 15-JUN-1990; 90US-00538857.  
PR 17-JUN-1991; 91US-00716725.  
PR 21-OCT-1994; 94US-00327381.  
XX  
PA (SCIO-) SCIOS INC.  
XX  
PI Cordell B;  
XX  
DR WPI; 1999-357231/30.  
XX  
PT Transgenic mice useful for studying compounds potentially useful in the

treatment of Alzheimer's disease.

Disclosure; Fig 8B; 72pp; English.

This invention describes novel transgenic mice expressing proteins related to the pathology of Alzheimer's disease and which provide models for studying potentially therapeutic compounds. The transgenic mice contain a DNA sequence encoding a beta-amyloid precursor protein (APP) and a nerve tissue specific promoter operably linked to the beta-APP allowing its expression to form beta-amyloid protein deposits in the animal's brain. The transgenic mouse is useful for elucidating the molecular mechanisms involved in the synthesis of and, more importantly, inhibiting the synthesis and deposition of beta-amyloid proteins (most importantly in the brain where plaque formation is associated with Alzheimer's disease) by inhibiting production and/or increasing cleavage after production. The transgenic animals provide useful models for studying the in vivo relationships of the proteins to each other and to other compounds being tested for their usefulness in treating Alzheimer's disease

Sequence 60 AA;

Query Match 78.7%; Score 37; DB 2; Length 60;  
Best Local Similarity 60.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GPCXXXFIRY 10  
||| |||  
14 GPCKAAMIRY 23

LT 14  
2769

AAW82769 standard; peptide; 128 AA.

AAW82769;

22-APR-1999 (first entry)

Bovine bikunin peptide variant #6.

Aprotinin; bovine; pancreatic trypsin inhibitor; BPTI; variant; bikunin; inhibitor; kallikrein; plasmin; factor Xa; elastase; blood loss; shock; thromboembolic disorder; polytrauma; sepsis; coagulation; organ failure; angina; myocardial infarct; stroke; embolism; thrombosis; restenosis; perfusion injury, inflammatory disorder; tumour invasion; metastasis; pain; oedema; haemostasis; dialysis.

Bos taurus.  
Synthetic.

DE19725014-A1.

17-DEC-1998.

13-JUN-1997; 97DE-01025014.

13-JUN-1997; 97DE-01025014.

(FARB ) BAYER AG.

Schroeder W, Apeler H;

WPI; 1999-046858/05.

Ne aprotinin and bikunin variants - useful for reducing blood loss during surgery, for treating e.g. thromboembolic disorders.

Disclosure; Fig 2; 40pp; German.

AAW82763-W82772 represent variants of the bovine bikunin peptide which is composed of two aprotinin (also known as bovine pancreatic trypsin inhibitor, BPTI) that have a net charge of +3 to -3 at pH 7. The

CC aprotinins are kallikrein, plasmin and factor Xa inhibitors. The bikunins  
CC are kallikrein, plasmin, factor Xa and elastase inhibitors. The  
CC aprotinins and bikunins can be used to reduce blood loss during surgery,  
CC to treat thromboembolic disorders, shock, polytrauma, sepsis,  
CC disseminated intravascular coagulation, multi-organ failure, unstable  
CC angina, myocardial infarct, stroke, embolism, deep vein thrombosis,  
CC restenosis, perfusion injury, thrombosis and bleeding after thrombosis,  
CC to treat inflammatory disorders, to inhibit tumour invasion and  
CC metastasis, to treat pain and oedema and to inhibit haemostasis  
CC activation during dialysis. The variants have altered or enhanced  
CC protease inhibition specificities

SQ Sequence 128 AA;

Query Match 78.7%; Score 37; DB 2; Length 128;  
Best Local Similarity 60.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| |||  
Db 82 GPCRAAIRY 91

RESULT 15

AAW82766

ID AAW82766 standard; peptide; 128 AA.

XX AAW82766;

XX 22-APR-1999 (first entry)

XX Bovine bikunin peptide, variant #3.

XX Aprotinin; bovine; pancreatic trypsin inhibitor; BPTI; variant; bikunin;  
XX inhibitor; kallikrein; plasmin; factor Xa; elastase; blood loss; shock;  
KW thromboembolic disorder; polytrauma; sepsis; coagulation; organ failure;  
KW angina; myocardial infarct; stroke; embolism; thrombosis; restenosis;  
KW perfusion injury, inflammatory disorder; tumour invasion; metastasis;  
KW pain; oedema; haemostasis; dialysis.

XX Bos taurus.  
OS Synthetic.

XX DE19725014-A1.

XX 17-DEC-1998.

XX 13-JUN-1997; 97DE-01025014.

XX 13-JUN-1997; 97DE-01025014.

XX (FARB ) BAYER AG.

XX Schroeder W, Apeler H;

XX WPI; 1999-046858/05.

XX Ne aprotinin and bikunin variants - useful for reducing blood loss during  
PT surgery, for treating e.g. thromboembolic disorders.

XX Disclosure; Fig 2; 40pp; German.

XX AAW82764-W82772 represent variants of the bovine bikunin peptide which is  
CC composed of two aprotinin (also known as bovine pancreatic trypsin  
CC inhibitor, BPTI) that have a net charge of +3 to -3 at pH 7. The  
CC aprotinins are kallikrein, plasmin and factor Xa inhibitors. The bikunins  
CC are kallikrein, plasmin, factor Xa and elastase inhibitors. The  
CC aprotinins and bikunins can be used to reduce blood loss during surgery,  
CC to treat thromboembolic disorders, shock, polytrauma, sepsis,  
CC disseminated intravascular coagulation, multi-organ failure, unstable  
CC angina, myocardial infarct, stroke, embolism, deep vein thrombosis,  
CC restenosis, perfusion injury, thrombosis and bleeding after thrombosis,  
CC to treat inflammatory disorders, to inhibit tumour invasion and

CC metastasis, to treat pain and oedema and to inhibit haemostasis  
CC activation during dialysis. The variants have altered or enhanced  
CC protease inhibition specificities

XX

SQ Sequence 128 AA;

Query Match 78.7%; Score 37; DB 2; Length 128;  
Best Local Similarity 60.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10

|||

Db 82 GPCRAAIRY 91

Search completed: August 11, 2004, 13:07:39  
Job time : 46.7143 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 11, 2004, 13:04:58 ; Search time 10.9524 Seconds  
(without alignments)  
87.827 Million cell updates/sec

le: US-10-661-207-1  
iect score: 47  
ience: 1 GPCXXXFIRY 10  
ing table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

ched: 283366 seqs, 96191526 residues

l number of hits satisfying chosen parameters: 283366

.um DB seq length: 0  
.um DB seq length: 2000000000  
rocessing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

ibase : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

ult no.	Score	Query		Length	DB	ID	Description
		Match					
1	42	89.4	122	1	A55115	uterine plasmin/tr	
2	39	83.0	57	2	A59204	basic proteinase i	
3	37	78.7	60	1	TIBOR	serum basic protei	
4	36	76.6	58	2	S10063	isoaprotinin G2 -	
5	36	76.6	59	2	S00371	isoaprotinin G1 -	
6	36	76.6	100	1	TIBO	basic proteinase i	
7	36	76.6	100	1	TIBOSP	spleen basic prote	
8	35	74.5	61	1	TIHCBP	proteinase inhibit	
9	35	74.5	62	2	S07451	proteinase inhibit	
10	35	74.5	67	1	TIBOC	trypsin inhibitor,	
11	35	74.5	324	2	AB0767	probable reductase	
12	35	74.5	330	1	S15303	probable CDP-6-deo	
13	35	74.5	580	2	T20716	hypothetical prote	
14	35	74.5	2225	2	T26063	hypothetical prote	
15	34	72.3	91	2	I68704	MHC c5/gII protein	
16	34	72.3	123	2	A29652	inter-alpha-trypsi	
17	34	72.3	235	2	A54951	tissue factor path	
18	34	72.3	250	2	I48850	TL antigen - mouse	
19	34	72.3	352	1	TIBOBI	alpha-1-microglobu	
20	34	72.3	361	2	B25132	MHC class I histoc	
21	34	72.3	361	2	I48851	TL antigen - mouse	
22	34	72.3	372	2	A25148	thymus leukemia an	
23	34	72.3	384	2	A25132	MHC class I histoc	
24	34	72.3	384	2	I54499	MHC thymus leukemi	
25	33	70.2	30	2	S07484	proteinase inhibit	
26	33	70.2	53	2	B29235	hemolymph trypsin	
27	33	70.2	1553	2	T09361	hypothetical prote	
28	32	68.1	56	2	JN0380	trypsin inhibitor	
29	32	68.1	58	1	TIHABK	isoinhibitor K (BP	

TL antigen - mouse  
alkaline exonuclea  
hepatocyte growth  
serine-repeat anti  
aspartate transami  
hypothetical prote  
serine-repeat anti  
SERA antigen/papai  
genome polypeptide  
hypothetical prote  
alanine-tRNA ligas  
C. perfringens tra  
hypothetical prote  
5'-amp-activated p  
SIP2 protein homol  
hypothetical prote

ALIGNMENTS

RESULT 1

A55115  
uterine plasmin/trypsin inhibitor - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A55115  
R;Stallings-Mann, M.L.; Burke, M.G.; Trout, W.E.; Roberts, R.M.  
J. Biol. Chem. 269, 24090-24094, 1994  
A;Title: Purification, characterization, and cDNA cloning of a Kunitz-type proteinase inh  
A;Reference number: A55115; MUID:95014140; PMID:7929061  
A;Accession: A55115  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-122 <STA>  
A;Cross-references: GB:I14282; NID:G682652; PIDN:AAA62425.1; PID:G682653  
A;Note: authors translated the codon GGC for residue 36 as Ala, AGC for residue 48 as Arg  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homolo  
C;Keywords: serine proteinase inhibitor  
F;38-88/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 89.4%; Score 42; DB 1; Length 122;  
Best local Similarity 60.0%; Pred. No. 0.071;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10

Db 45 GPCSAHFVRY 54

RESULT 2

A59204  
basic proteinase inhibitor - great pond snail  
N;Alternate names: trypsin inhibitor  
C;Species: Lymnaea stagnalis (great pond snail)  
C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 17-Mar-2000  
C;Accession: A59204  
R;Nagle, G.  
submitted to the Protein Sequence Database, March 2000  
A;Description: Lymnaea trypsin inhibitor.  
A;Reference number: A59204  
A;Accession: A59204  
A;Molecule type: protein  
A;Residues: 1-57 <NAG>  
A;Experimental source: albumen gland  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homolo  
C;Keywords: glycoprotein; serine proteinase inhibitor  
F;5-55/Domain: animal Kunitz-type proteinase inhibitor  
F;5-55,14-38,30-51/Disulfide bonds: #status predicted  
F;15/inhibitory site: Lys (trypsin) #status predicted  
F;24/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 83.0%; Score 39; DB 2; Length 57;

Best Local Similarity 60.0%; Pred. No. 0.16;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| |:  
Db 12 GPCKGNFLRY 21

RESULT 3

TIBOR  
serum basic proteinase inhibitor - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 31-Oct-1980 #sequence\_revision 31-Oct-1980 #text\_change 31-Dec-1993  
C;Accession: A01206  
R;Wachter, E.; Deppner, K.; Hochstrasser, K.; Lempart, K.; Geiger, R.  
FEBS Lett. 119, 58-62, 1980  
A;Title: A new Kunitz-type inhibitor from bovine serum amino acid sequence determination  
A;Reference number: A01206; MUID:81044408; PMID:7428928  
A;Accession: A01206

A;Molecule type: protein  
A;Residues: 1-60 <WAC>  
C;Comment: This inhibitor has activity very similar to that of the basic proteinase inhibitor  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homol  
C;Keywords: serine proteinase inhibitor  
F;7-57/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F;17/Inhibitory site: Lys (trypsin) #status experimental

Query Match 78.7%; Score 37; DB 1; Length 60;  
Best Local Similarity 60.0%; Pred. No. 0.46;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| |:  
Db 14 GPCKAAMIRY 23

RESULT 4

S10063  
isoaprotinin G2 - bovine hybrid  
C;Species: Bos primigenius indicus x Bos primigenius taurus (cattle)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 13-Jun-1997  
C;Accession: S10063  
R;Siekmann, J.; Wenzel, H.R.; Schroeder, W.; Tschesche, H.  
Biol. Chem. Hoppe-Seyler 369, 157-163, 1988  
A;Title: Characterization and sequence determination of six aprotinin homologues from bo  
A;Reference number: S00371; MUID:88221840; PMID:2453200

A;Accession: S10063  
A;Molecule type: protein  
A;Residues: 1-58 <SIE>  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homol  
C;Keywords: serine proteinase inhibitor  
F;5-55/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F;15/Inhibitory site: Lys (trypsin) #status predicted

Query Match 76.6%; Score 36; DB 2; Length 58;  
Best Local Similarity 60.0%; Pred. No. 0.73;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| |:  
Db 12 GPCKARMIRY 21

RESULT 5

S00371  
isoaprotinin G1 - bovine hybrid  
C;Species: Bos primigenius indicus x Bos primigenius taurus (cattle)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 13-Jun-1997  
C;Accession: S00371  
R;Siekmann, J.; Wenzel, H.R.; Schroeder, W.; Tschesche, H.  
Biol. Chem. Hoppe-Seyler 369, 157-163, 1988  
A;Title: Characterization and sequence determination of six aprotinin homologues from bo  
A;Reference number: S00371; MUID:88221840; PMID:2453200

A;Accession: S00371  
A;Molecule type: protein  
A;Residues: 1-59 <SIE>

C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor hor  
C;Keywords: pyroglutamic acid; serine proteinase inhibitor  
F;6-56/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
F;16/Inhibitory site: Lys (trypsin) #status predicted

Query Match 76.6%; Score 36; DB 2; Length 59;  
Best Local Similarity 60.0%; Pred. No. 0.74;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| |:  
Db 13 GPCKARMIRY 22

RESULT 6

TIBO  
basic proteinase inhibitor precursor - bovine  
N;Alternate names: aprotinin; basic pancreatic trypsin inhibitor; BPTI; cationic kall  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 24-Apr-1984 #sequence\_revision 22-Jul-1994 #text\_change 16-Jun-2000  
C;Accession: S00277; A30333; S10546; S02486; S28197; A90162; A92023; A90736; A90927; J  
R;Creighton, T.E.; Charles, I.G.  
J. Mol. Biol. 194, 11-22, 1987

A;Title: Sequences of the genes and polypeptide precursors for two bovine protease in  
A;Reference number: S00274; MUID:87283904; PMID:2441071  
A;Accession: S00277

A;Molecule type: DNA; mRNA  
A;Residues: 1-100 <CR2>

A;Cross-references: GB:M20934; GB:X05274; NID:gl62767; PIDN:AAD13685.1; PID:gl62769  
R;Creighton, T.E.; Charles, I.G.  
Cold Spring Harb. Symp. Quant. Biol. 52, 511-519, 1987

A;Title: Biosynthesis, processing, and evolution of bovine pancreatic trypsin inhibit  
A;Reference number: A90926; MUID:88295740; PMID:2456884

A;Accession: A30333  
A;Molecule type: DNA

A;Residues: 1-100 <CRE>

A;Cross-references: GB:M20934; GB:X05274; NID:gl62767; PIDN:AAD13685.1; PID:gl62769  
R;Kingston, I.B.; Anderson, S.  
Biochem. J. 233, 443-450, 1986

A;Title: Sequences encoding two trypsin inhibitors occur in strikingly similar genom  
A;Reference number: S10546; MUID:86158754; PMID:2420326

A;Accession: S10546  
A;Molecule type: DNA

A;Residues: 34-97 <KIN>  
R;Fioretti, E.; Angeletti, M.; Fiorucci, L.; Barra, D.; Bossa, F.; Ascoli, F.  
Biol. Chem. Hoppe-Seyler 369(Suppl.), 37-42, 1988

A;Title: Aprotinin-like isoinhibitors in bovine organs.

A;Reference number: S02485; MUID:89076531; PMID:2462435

A;Accession: S02486  
A;Molecule type: protein

A;Residues: 36-93 <FIO>  
R;Ikekita, M.; Jone, C.S.; Kamo, M.; Tsugita, A.; Kizuki, K.; Moriya, H.  
Protein Seq. Data Anal. 5, 7-11, 1992

A;Title: Purification and characterization of the major cationic kallikrein inhibitor  
A;Reference number: S28197; MUID:93150003; PMID:1283464

A;Accession: S28197  
A;Molecule type: protein

A;Residues: 36-93 <IKE>  
R;Kassell, B.; Laskowski, M.

Biochem. Biophys. Res. Commun. 20, 463-468, 1965

A;Title: The basic trypsin inhibitor of bovine pancreas. V. The disulfide linkages.

A;Reference number: A90162; MUID:66083012; PMID:5860161

A;Contents: annotation; disulfide bonds

A;Accession: A90162  
A;Molecule type: protein

A;Residues: 36-93 <KAS>  
R;Anderer, F.A.; Hornie, S.

J. Biol. Chem. 241, 1568-1572, 1966  
A;Title: The disulfide linkages in kallikrein inactivator of bovine lung.

Reference number: A92023; MUID:66171231; PMID:5296424  
 Content: annotation; disulfide bonds  
 Accession: A92023  
 Molecule type: protein  
 Residues: 36-93 <AN2>  
 Lauvet, J.; Acher, R.  
 Soc. Chim. Biol. 49, 985-1000, 1967  
 Title: La structure covalente d'un inhibiteur polypeptidique de la trypsine (inhibited)  
 Reference number: A90736; MUID:68012003; PMID:6053284  
 Content: annotation; disulfide bonds  
 Accession: A90736  
 Molecule type: protein  
 Residues: 36-93 <CHA>  
 ouha, V.; Pospisilova, D.; Meloun, B.; Sorm, F.  
 Czech. Chem. Commun. 33, 1363-1365, 1968  
 Title: Sequence of residues 18-20 in pancreatic trypsin inhibitor.  
 Reference number: A90927  
 Accession: A90927  
 Molecule type: protein  
 Residues: 36-93 <DLO>  
 ber, R.; Kukla, D.; Ruhlmann, A.; Epp, O.; Formanek, H.  
 wissenshaften 57, 389-392, 1970  
 Title: The basic trypsin inhibitor of bovine pancreas. I. Structure analysis and conf  
 Reference number: A93410; MUID:70255230; PMID:5447861  
 Content: annotation; X-ray crystallography of basic protease inhibitor, 2.5 angstroms  
 wis, R.V.; Ray, P.; Cogull, R.; Kruggel, W.  
 hem. Biophys. Res. Commun. 167, 543-547, 1990  
 Title: Presence of pancreatic trypsin inhibitor in adrenal medullary chromaffin cells.  
 Reference number: A34658; MUID:90211226; PMID:2322242  
 Accession: A34658  
 Molecule type: protein  
 Residues: 36-53, 55-81 <LEW>  
 erson, S.; Kingston, I.B.  
 . Natl. Acad. Sci. U.S.A. 80, 6838-6842, 1983  
 Title: Isolation of a genomic clone for bovine pancreatic trypsin inhibitor by using a  
 Reference number: A93977; MUID:84070725; PMID:6580617  
 Accession: A93977  
 Molecule type: DNA  
 Residues: 'PSLFNRDPPIPA', 34-97, 'GKTGGRAEGEGKG' <AND>  
 oss-references: GB:X03365; GB:X00966; NID:gl142; PIDN:CAA27062.1; PID:gl1364183  
 ekmann, J.; Wenzel, H.R.; Schroeder, W.; Tschesche, H.  
 Chem. Hoppe-Seyler 369, 157-163, 1988  
 Title: Characterization and sequence determination of six aprotinin homologues from bo  
 Reference number: S00371; MUID:88221840; PMID:2453200  
 Accession: S10062  
 Molecule type: protein  
 Residues: 36-66, 'P', 68-82, 'S', 84-93 <SIE>  
 perimental source: lung  
 Note: the authors designated this protein as isoaprotinin 2  
 Content: Basic protease inhibitor is an intracellular polypeptide found in many tiss  
 netics:  
 Introns: 34/1; 98/1  
 perfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homol  
 ywords: serine proteinase inhibitor  
 20/Domain: signal sequence #status predicted <SIG>  
 .-35/Domain: propeptide #status predicted <PRO>  
 .-100/Product: basic proteinase inhibitor #status experimental <MAT>  
 .-90/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
 .-90, 49-73, 65-86/Disulfide bonds: #status experimental  
 /Inhibitory site: Lys (trypsin, chymotrypsin, kallikrein, plasmin) #status experimen  
 6ery Match 76.6%; Score 36; DB 1; Length 100;  
 1st Local Similarity 60.0%; Pred. No. 1.2;  
 itches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 1 GPCXXXFYRY 10  
 |||||  
 47 GPCKARIIRY 56  
 LT 7  
 SP  
 en basic proteinase inhibitor precursor - bovine

N;Alternate names: isoaprotinin 1  
 C;Species: Bos primigenius taurus (catle)  
 C;Date: 07-Sep-1990 #sequence\_revision 22-Jul-1994 #text\_change 21-Jul-2000  
 C;Accession: S00274; B30333; A27417; S02487; S10547; S13478; A23915; S10064  
 R;Creighton, T.E.; Charles, I.G.  
 J. Mol. Biol. 194, 11-22, 1987  
 A;Title: Sequences of the genes and polypeptide precursors for two bovine protease inhib  
 A;Reference number: S00274; MUID:87283904; PMID:2441071  
 A;Accession: S00274  
 A;Molecule type: DNA; mRNA  
 A;Residues: 1-100 <CR2>  
 A;Cross-references: EMBL:X05275; NID:gl163718; PIDN:AAA51418.1; PID:gl163720  
 R;Creighton, T.E.; Charles, I.G.  
 Cold Spring Harb. Symp. Quant. Biol. 52, 511-519, 1987  
 A;Title: Biosynthesis, processing, and evolution of bovine pancreatic trypsin inhibitor.  
 A;Reference number: A90926; MUID:88295740; PMID:2456884  
 A;Accession: B30333  
 A;Molecule type: DNA  
 A;Residues: 1-100 <CRE>  
 A;Cross-references: GB:M20935; GB:X05275; NID:gl163718; PIDN:AAA51418.1; PID:gl163720  
 R;Barra, D.; Simmaco, M.; Bossa, F.; Fioretti, E.; Angeletti, M.; Ascoli, F.  
 J. Biol. Chem. 262, 13916-13919, 1987  
 A;Title: Primary structure of a protease isoinhibitor from bovine spleen. A possible inte  
 A;Reference number: A27417; MUID:88007630; PMID:3654647  
 A;Accession: A27417  
 A;Molecule type: protein  
 A;Residues: 34-99 <BAR>  
 A;Note: the more abundant form of isoinhibitor I lacks 99-Asn  
 R;Fioretti, E.; Angeletti, M.; Fiorucci, L.; Barra, D.; Bossa, F.; Ascoli, F.  
 Biol. Chem. Hoppe-Seyler 369 (Suppl.), 37-42, 1988  
 A;Title: Aprotinin-like isoinhibitors in bovine organs.  
 A;Reference number: S02485; MUID:89076531; PMID:2462435  
 A;Accession: S02487  
 A;Molecule type: protein  
 A;Residues: 34-99 <FIO>  
 R;Kingston, I.B.; Anderson, S.  
 Biochem. J. 233, 443-450, 1986  
 A;Title: Sequences encoding two trypsin inhibitors occur in strikingly similar genomic er  
 A;Reference number: S10546; MUID:86158754; PMID:2420326  
 A;Accession: S10547  
 A;Molecule type: DNA  
 A;Residues: 34-97 <KIN>  
 R;Barra, D.; Fioretti, E.; Angeletti, M.; Maras, B.; Bossa, F.; Ascoli, F.  
 Biochim. Biophys. Acta 1076, 143-147, 1991  
 A;Title: Proteinase isoinhibitors from bovine spleen: primary structure of an intermediat  
 A;Reference number: S13478; MUID:91098258; PMID:1986787  
 A;Accession: S13478  
 A;Molecule type: protein  
 A;Residues: 36-97 <BA2>  
 R;Fioretti, E.; Iacopino, G.; Angeletti, M.; Barra, D.; Bossa, F.; Ascoli, F.  
 J. Biol. Chem. 260, 11451-11455, 1985  
 A;Reference number: A23915; MUID:86008178; PMID:2413011  
 A;Accession: A23915  
 A;Molecule type: protein  
 A;Residues: 36-93 <FI2>  
 A;Note: three disulfide bonds are present  
 R;Siekmann, J.; Wenzel, H.R.; Schroeder, W.; Tschesche, H.  
 Biol. Chem. Hoppe-Seyler 369, 157-163, 1988  
 A;Title: Characterization and sequence determination of six aprotinin homologues from bo  
 A;Reference number: S00371; MUID:88221840; PMID:2453200  
 A;Accession: S10064  
 A;Molecule type: protein  
 A;Residues: 36-93 <SIE>  
 A;Experimental source: lung  
 C;Comment: This inhibitor is found in many bovine organs. In spleen, it is located in var  
 C;Comment: Compared with the basic proteinase inhibitor, affinity for trypsin and chymotr  
 C;Comment: Inhibitors I and III are intermediates in the processing of the precursor to t  
 C;Genetics:  
 A;Introns: 34/1; 98/1  
 C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homol  
 C;Keywords: serine proteinase inhibitor  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-33/Domain: propeptide #status predicted <PRO>



F;34-98/Product: spleen inhibitor I #status experimental <SI1>  
F;36-97/Product: spleen inhibitor III #status experimental <SI3>  
F;36-93/Product: spleen inhibitor II #status experimental <SI2>  
F;40-90/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F;40-90,49-73,65-86/Disulfide bonds: #status predicted  
F;50/Inhibitory site: Lys (trypsin, chymotrypsin, kallikrein) #status predicted

Query Match 76.6%; Score 36; DB 1; Length 100;  
Best Local Similarity 60.0%; Pred. No. 1.2;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| |||  
Db 47 GPCAKMIRY 56

## RESULT 8

TIHCBP

Proteinase inhibitor (BPI-type) - horseshoe crab (Tachypleus tridentatus)  
C;Species: Tachypleus tridentatus  
C;Date: 08-Mar-1989 #sequence\_revision 22-Jul-1994 #text\_change 24-Feb-1995  
C;Accession: A26923  
R;Nakamura, T.; Hirai, T.; Tokunaga, F.; Kawabata, S.; Iwanaga, S.  
J. Biochem. 101, 1297-1306, 1987  
A;Title: Purification and amino acid sequence of Kunitz-type protease inhibitor found in  
A;Reference number: A26923; MUID:88007472; PMID:3308864  
A;Accession: A26923  
A;Molecule type: protein  
A;Residues: 1-61 <NAK>  
A;Experimental source: hemocytes  
C;Comment: The inhibitory activity is similar to bovine basic proteinase inhibitor.  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homol  
C;Keywords: serine proteinase inhibitor  
F;8-58/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F;8-58,17-41,33-54/Disulfide bonds: #status predicted  
F;18/Inhibitory site: Arg (chymotrypsin, elastase, trypsin, plasmin, plasma kallikrein)

Query Match 74.5%; Score 35; DB 1; Length 61;  
Best Local Similarity 60.0%; Pred. No. 1.3;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| |||  
Db 15 GPCRAGFKRY 24

## RESULT 9

S07451

Proteinase inhibitor 5.II - snake-locks sea anemone  
C;Species: Anemonia sulcata (snake-locks sea anemone)  
C;Date: 31-Dec-1990 #sequence\_revision 09-May-1997 #text\_change 09-May-1997  
C;Accession: S07451; B27222  
R;Wunderer, G.; Machleidt, W.; Fritz, H.  
Meth. Enzymol. 80, 816-820, 1981  
A;Title: The broad-specificity proteinase inhibitor 5 II from the sea anemone Anemonia s  
A;Reference number: S07451  
A;Accession: S07451  
A;Molecule type: protein  
A;Residues: 1-59 <WUN>  
A;Note: 13-Arg, 16-Gly, 17-Gly, 25-Leu, 28-Arg, and 39-Arg were also found  
R;Krebs, H.C.; Habermehl, G.G.  
Naturwissenschaften 74, 395-396, 1987  
A;Title: Isolierung und strukturaufklaerung eines haemolytisch aktiven peptids aus der s  
A;Reference number: A94700  
A;Accession: B27222  
A;Molecule type: protein  
A;Residues: 1-38, 'R', '40', 'B', '42', 'BB', '45-48', 'ZZ', '51', 'Z', '53-62 <KRE>  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homol  
C;Keywords: serine proteinase inhibitor  
F;5-55/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 74.5%; Score 35; DB 2; Length 62;  
Best Local Similarity 60.0%; Pred. No. 1.3;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GPCXXXFIRY 10  
||| |||  
Db 12 GPCRARFPY 21

## RESULT 10

TIBOC

trypsin inhibitor, colostrum (BPI type) - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 24-Apr-1984 #sequence\_revision 31-Dec-1993 #text\_change 06-Dec-1996  
C;Accession: A01207  
R;Cechova, D.; Jonakova, V.; Sorm, F.  
Collect. Czech. Chem. Commun. 36, 3342-3357, 1971  
A;Title: Primary structure of trypsin inhibitor from cow colostrum (component B2).  
A;Reference number: A90928  
A;Accession: A01207  
A;Molecule type: protein  
A;Residues: 1-26, 'B', 28-67 <CEC>  
A;Note: the residue identified as Asx is bound to carbohydrate; therefore, we have shc  
R;Cechova, D.; Ber, E.  
Collect. Czech. Chem. Commun. 39, 680-688, 1974  
A;Title: Disulfide bonds of trypsin inhibitor from cow colostrum.  
A;Reference number: A90929  
A;Contents: annotation; disulfide bonds  
R;Cechova, D.; Muszynska, G.  
FEBS Lett. 8, 84-86, 1970  
A;Title: Role of lysine 18 in active center of cow colostrum trypsin inhibitor.  
A;Reference number: A91440  
A;Contents: annotation

A;Note: the inhibitory site was determined  
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor hor  
C;Keywords: colostrum; glycoprotein; serine proteinase inhibitor  
F;8-58/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F;8-58,17-41,33-54/Disulfide bonds: #status experimental  
F;18/Inhibitory site: Lys (trypsin) #status experimental  
F;27/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 74.5%; Score 35; DB 1; Length 67;  
Best Local Similarity 50.0%; Pred. No. 1.4;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| |||  
Db 15 GPCKAALLRY 24

## RESULT 11

AB0767

probable reductase RfBI [imported] - Salmonella enterica subsp. enterica serovar  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AB0767  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
ch, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, I  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AB0767  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-324 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD02456.1; PID:g16503323; GSPDB:GN00176  
C;Genetics:  
A;Gene: rfBI  
C;Superfamily: methane monooxygenase reductase component; cytochrome-b5 reductase hom

Query Match 74.5%; Score 35; DB 2; Length 324;  
Best Local Similarity 66.7%; Pred. No. 5.4;



atches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
1 GPCXXXFIR 9  
184 GPCGTFIR 192  
JLT 12  
03  
able CDP-6-deoxy-Delta(3,4)-glucose reductase (EC 1.3.1.-) - Salmonella typhimurium  
ternate names: hypothetical protein 7.6  
pecies: Salmonella typhimurium  
te: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
ession: S15303  
ang, X.M.; Neal, B.; Santiago, F.; Lee, S.J.; Romana, L.K.; Reeves, P.R.  
Microbiol. 5, 695-713, 1991  
tle: Structure and sequence of the rfb (O antigen) gene cluster of Salmonella serovar  
ference number: S15296; MUID:91260454; PMID:1710759  
ession: S15303  
atus: preliminary  
lecle type: DNA  
sides: 1-330 <MOL>  
perfamily: methane monooxygenase reductase component; cytochrome-b5 reductase homolog  
ywords: 2Fe-2S; iron-sulfur protein; metalloprotein; oxidoreductase  
-72/Domain: ferredoxin [2Fe-2S] homology <FER>  
1-320/Domain: cytochrome-b5 reductase homology <CBR>  
, 42,45,71/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted  
ery Match 74.5%; Score 35; DB 1; Length 330;  
st Local Similarity 66.7%; Pred. No. 5.4;  
atches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
1 GPCXXXFIR 9  
184 GPCGTFIR 192  
JLT 13  
16  
thetical protein F10F2.8 - Caenorhabditis elegans  
pecies: Caenorhabditis elegans  
te: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
ession: T20716  
les, L.  
itted to the EMBL Data Library, August 1994  
ference number: Z19313  
atus: preliminary; translated from GB/EMBL/DBJ  
lecle type: DNA  
sides: 1-580 <WIL>  
ross-references: EMBL:Z35598; PIDN:CAA84654.1; GSPDB:GN00021; CESP:F10F2.8  
perimental source: clone F10F2  
netics:  
ne: CESP:F10F2.8  
p position: 3  
itrons: 46/3; 64/3; 151/1; 330/3; 365/2; 453/1  
ery Match 74.5%; Score 35; DB 2; Length 580;  
st Local Similarity 50.0%; Pred. No. 8.9;  
atches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
1 GPCXXXFIR 10  
395 GPCPAGFLQY 404  
JLT 14  
63  
thetical protein W01F3.3 - Caenorhabditis elegans  
pecies: Caenorhabditis elegans  
te: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
ession: T26063  
mmings, P.

submitted to the EMBL Data Library, March 1997  
A;Reference number: Z20145  
A;Accession: T26063  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-2225 <WIL>  
A;Cross-references: EMBL:Z92815; PIDN:CAB07294.1; GSPDB:GN00023; CESP:W01F3.3  
A;Experimental source: clone W01F3  
C;Genetics:  
A;Gene: CESP:W01F3.3  
A;Map position: 5  
A;Introns: 33/1; 56/1; 100/1; 142/3; 271/3; 451/1; 525/3; 774/1; 1093/1; 1178/1; 1221/1;  
Query Match 74.5%; Score 35; DB 2; Length 2225;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GPCXXXFIR 10  
Db 470 GPCGCSFQRY 479  
RESULT 15  
I68704  
MHC c5/gll protein - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C;Accession: I68704  
R;Rogers, J.H.  
Immunogenetics 21, 343-353, 1985  
A;Title: Family organization of mouse H-2 class I genes.  
A;Reference number: I54413; MUID:85206117; PMID:3997208  
A;Accession: I68704  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-91 <RES>  
A;Cross-references: GB:M14828; NID:G199643; PIDN:AAA39688.1; PID:G554240  
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology  
Query Match 72.3%; Score 34; DB 2; Length 91;  
Best Local Similarity 50.0%; Pred. No. 2.9;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GPCXXXFIR 10  
Db 71 GPCKDSLRY 80  
Search completed: August 11, 2004, 13:09:58  
Job time : 10.9524 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 11, 2004, 13:01:12 ; Search time 7.61905 Seconds  
(without alignments)  
68.342 Million cell updates/sec

le: US-10-661-207-1  
ect score: 47  
ience: 1 GPCXXFXFIRY 10

ing table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

ched: 141681 seqs, 52070155 residues

al number of hits satisfying chosen parameters: 141681

imum DB seq length: 0  
imum DB seq length: 2000000000

:-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

ibase : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	DB	ID	Description
1	42	89.4	122	1	UPTI_PIG	Q29100 sus scrofa
2	37	78.7	25	1	BT14_BOOMI	P83605 boophilus m
3	37	78.7	60	1	IBPS_BOVIN	P00975 bos taurus
4	37	78.7	133	1	EPPI_MACMU	Q9bd11 macaca mula
5	36	76.6	100	1	BPT1_BOVIN	P00974 bos taurus
6	36	76.6	100	1	BPT2_BOVIN	P04815 bos taurus
7	35	74.5	55	1	CSTI_BOMMO	P81902 bombyx mori
8	35	74.5	61	1	IBPI_TACTR	P16044 tachypleus
9	35	74.5	62	1	IP52_ANESU	P10280 anemonia su
10	35	74.5	67	1	IBPC_BOVIN	P00976 bos taurus
11	35	74.5	197	1	MCPI_MELCP	P82968 melithaea c
12	35	74.5	330	1	RFBI_SALTY	P26395 salmonella
13	34	72.3	123	1	IATR_SHEEP	P13371 ovis aries
14	34	72.3	235	1	TFP2_HUMAN	P48307 homo sapien
15	34	72.3	352	1	AMBP_BOVIN	P00978 bos taurus
16	34	72.3	361	1	HA1U_MOUSE	P14433 mus musculu
17	34	72.3	384	1	HA1T_MOUSE	P14432 mus musculu
18	33	70.2	53	1	HTIB_MANSE	P26227 manduca sex
19	32	68.1	56	1	ITR4_RADMA	P16344 radianthus
20	32	68.1	58	1	ISIK_HELPO	P00994 helix pomat
21	32	68.1	470	1	EXON_EBV	P03217 epstein-bar
22	32	68.1	883	1	SYA_EUCAP	Q8k9e7 buchnera ap
23	32	68.1	1544	1	TUL4_HUMAN	Q9nrj4 homo sapien
24	32	68.1	1547	1	TUL4_MOUSE	Q9j1l5 mus musculu
25	31	66.0	252	1	SPT2_MOUSE	O43291 homo sapien
26	31	66.0	252	1	SPT2_MOUSE	Q9wu03 mus musculu
27	31	66.0	370	1	AAT2_METJA	Q58097 methanococc
28	31	66.0	701	1	ACSA_HUMAN	Q9nr19 homo sapien
29	31	66.0	989	1	SERA_PLAFG	P13823 plasmodium
30	31	66.0	2127	1	RRPL_RABVS	P16289 rabies viru
31	30	63.8	133	1	EPPI_HUMAN	Q95925 homo sapien
32	30	63.8	201	1	SYA_RHILV	P24075 rhizobium l
33	30	63.8	310	1	OK4E_HUMAN	Q8ngd5 homo sapien

ALIGNMENTS

RESULT 1

UPTI\_PIG

ID UPTI\_PIG STANDARD; PRT; 122 AA.

AC Q29100;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Uterine plasmin/trypsin inhibitor precursor (UPTI).

OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]

RP SEQUENCE FROM N.A., SEQUENCE OF 30-56, AND CHARACTERIZATION.  
RC TISSUE=Conceptus membrane, and Uterus;  
RX MEDLINE=95014140; PubMed=7929061;  
RA Stallings-Mann M.L., Burke M.G., Trout W.E., Roberts R.M.;  
RT "Purification, characterization, and cDNA cloning of a Kunitz-type  
RT proteinase inhibitor secreted by the porcine uterus.";  
RL J. Biol. Chem. 269:24090-24094(1994).

CC -I- FUNCTION: Inhibitor of plasmin and trypsin. Also has a weak  
CC affinity for chymotrypsin. Colud serves to neutralize the  
CC activities of one or more serine proteinases generated by the  
CC proliferating trophoblast during the formation of the noninvasive  
CC placenta.

CC -I- TISSUE SPECIFICITY: Expressed only in the uterus and the  
CC endometrium.

CC -I- DEVELOPMENTAL STAGE: Maximally expressed during pregnancy until  
CC day 30 after which levels decrease significantly.

CC -I- INDUCTION: By progesterone.

CC -I- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; L14282; AA62425.1; -.

DR PIR; A55115; A55115.

DR HSSP; P00974; IBPI.

DR InterPro; IPR002223; Kunitz\_BPTI.

DR Pfam; PF00014; Kunitz\_BPTI\_1.

DR PRINTS; PR00759; BASICTPASE.

DR ProDom; PD000222; Kunitz\_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.

KW Serine protease inhibitor; Signal.

FT SIGNAL 1 20 POTENTIAL.

FT PROPEP 21 29

FT CHAIN 30 122

FT DOMAIN 38 88

FT DISULFID 38 88

UTERINE PLASMIN/TRYPsin INHIBITOR.

BPTI/KUNITZ INHIBITOR.

BY SIMILARITY.

FT DISULFID 47 71 BY SIMILARITY.  
FT ACT SITE 63 84 BY SIMILARITY.  
FT CONFLICT 48 49 REACTIVE BOND (BY SIMILARITY).  
FT CONFLICT 36 36 G -> A (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 48 48 S -> R (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 52 52 V -> I (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 54 54 Y -> V (IN REF. 1; AA SEQUENCE).  
SQ SEQUENCE 122 AA; 13109 MW; B9BFB3A6FB76CEC CRC64;

Query Match 89.4%; Score 42; DB 1; Length 122;  
Best Local Similarity 60.0%; Pred. No. 0.019;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |||  
DB 45 GPCSAHFVRY 54

## RESULT 2

BT14 BOOMI STANDARD; PRT; 25 AA.  
ID BT14 BOOMI STANDARD; PRT; 25 AA.  
AC P83605;

DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Kunitz-type serine protease inhibitor BmtI-4 (Fragment).

OS Boophilus microplus (Cattle tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.

NCBI\_TaxID=6941;

## SEQUENCE, AND FUNCTION.

TISSUE=Larva;

Sasaki S.D., Hirata I.Y., Tanaka A.S.;

"Molecular studies of serine protease inhibitors from cattle tick  
Boophilus microplus (larvae).";  
Submitted (JUN-2003) to Swiss-Prot.

-!- FUNCTION: Inhibits trypsin, human plasma kallikrein and human  
neutrophil elastase.

-!- SUBCELLULAR LOCATION: Secreted.

-!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

InterPro; IPR002223; Kunitz\_BPTI.

PROSITE; PS00280; BPTI\_KUNITZ\_1; PARTIAL.

PROSITE; PSS0279; BPTI\_KUNITZ\_2; PARTIAL.

Serine protease inhibitor.

ACT SITE 18 19 REACTIVE BOND (BY SIMILARITY).

NON\_TER 25 25

SEQUENCE 25 AA; 2963 MW; 81CB3A2D2E121F3D CRC64;

## Query Match

Best Local Similarity 78.7%; Score 37; DB 1; Length 25;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |||  
DB 15 GPCVGYFFRY 24

## RESULT 3

IBPS BOVIN STANDARD; PRT; 60 AA.  
ID IBPS BOVIN STANDARD; PRT; 60 AA.  
AC P00975;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Serum basic protease inhibitor (Serum BPI).

OS Bos taurus (Bovine).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

NCBI\_TaxID=9913;

## SEQUENCE.

RX MEDLINE=81044408; PubMed=7428928;  
RA Wachter E., Deppner K., Hochstrasser K., Lempart K., Geiger R.;  
RT "A new Kunitz-type inhibitor from bovine serum amino acid sequence  
determination.";  
RL FEBS Lett. 119:58-62(1980).  
CC -!- FUNCTION: This inhibitor has activity very similar to that of the  
basic protease inhibitor from bovine tissues.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

PIR; A01206; TIBOR.

HSSP; P00974; 1BPI.

InterPro; IPR002223; Kunitz\_BPTI.

Pfam; PF00014; Kunitz\_BPTI\_1.

PRINTS; PR00759; BASICPTASE.

ProDom; PD000222; Kunitz\_BPTI; 1.

SMART; SM00131; KU; 1.

PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

PROSITE; PSS0279; BPTI\_KUNITZ\_2; 1.

Serine protease inhibitor.

DISULFID 7 57 BY SIMILARITY.

DISULFID 16 40 BY SIMILARITY.

DISULFID 32 53 BY SIMILARITY.

ACT SITE 17 18 REACTIVE BOND.

SEQUENCE 60 AA; 6647 MW; B9953EBAACFIA4E6 CRC64;

Query Match 78.7%; Score 37; DB 1; Length 60;

Best Local Similarity 60.0%; Pred. No. 0.12;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |||

DB 14 GPCKAAMIRY 23

## RESULT 4

EPPI\_MACMU

ID EPPI\_MACMU STANDARD; PRT; 133 AA.

AC Q9BDL1;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Eppin precursor (Epididymal protease inhibitor) (Serine protease

inhibitor-like with Kunitz and WAP domains 1).

GN SPINLW1.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

OC Cercopitheidae; Macaca.

NCBI\_TaxID=9544;

## SEQUENCE FROM N.A.

TISSUE=Epididymis, and Testis;

Sivashanmugam P., Hall S.H., Hamil K.G., French F.S., O'Rand M.G.,

Richardson R.T.;

"Characterization of monkey and mouse Eppin, a protease inhibitor from

epididymis and testis.";

Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

-!- SUBCELLULAR LOCATION: Secreted (Potential).

-!- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.

-!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

-!- SIMILARITY: Contains 1 WAP-type domain.

## This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF346414; AAK31336.1; -

HSSP; P00974; 1BPI.

InterPro; IPR002223; Kunitz\_BPTI.



InterPro; IPR008197; WAP.  
PFam; PF00014; Kunitz\_BPTI; 1.  
PFam; PF00095; wap; 1.  
PRINTS; PR00759; BASICPTASE.  
ProDom; PD000222; Kunitz\_BPTI; 1.  
SMART; SM00131; KU; 1.  
SMART; SM00217; WAP; 1.  
PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
PROSITE; PS0279; BPTI\_KUNITZ\_2; 1.  
Serine protease inhibitor; Signal.  
SIGNAL 1 21 POTENTIAL.  
CHAIN 22 133 EPPIN.  
DOMAIN 77 127 WAP.  
DOMAIN 77 127 BPTI/KUNITZ INHIBITOR.  
DISULFID 33 61 BY SIMILARITY.  
DISULFID 40 65 BY SIMILARITY.  
DISULFID 48 60 BY SIMILARITY.  
DISULFID 54 69 BY SIMILARITY.  
DISULFID 77 127 BY SIMILARITY.  
DISULFID 86 110 BY SIMILARITY.  
DISULFID 102 123 BY SIMILARITY.  
SEQUENCE 133 AA; 15279 MW; 433AE946E39A35E9 CRC64;  
very Match 78.7%; Score 37; DB 1; Length 133;  
st Local Similarity 60.0%; Pred. No. 0.27;  
atches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
1 GPCXXXFIRY 10  
||| |||  
84 GPCLAFFIRW 93  
ILT 5  
\_BOVIN STANDARD; PRT; 100 AA.  
P00974;  
21-JUL-1986 (Rel. 01, Created)  
01-MAR-1989 (Rel. 10, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Pancreatic trypsin inhibitor precursor (Basic protease inhibitor)  
(BPI) (BPTI) (Aprotinin).  
Bos taurus (Bovine).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
Bovidae; Bovinae; Bos.  
NCBI\_TaxID=9913;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=87283904; PubMed=2441071;  
Creighton T.E., Charles I.G.;  
"Sequences of the genes and polypeptide precursors for two bovine  
protease inhibitors.";  
J. Mol. Biol. 194:11-22(1987).  
[2]  
SEQUENCE FROM N.A.  
MEDLINE=88295740; PubMed=2456884;  
Creighton T.E., Charles I.G.;  
"Biosynthesis, processing, and evolution of bovine pancreatic trypsin  
inhibitor.";  
Cold Spring Harb. Symp. Quant. Biol. 52:511-519(1987).  
[3]  
SEQUENCE OF 34-97 FROM N.A.  
MEDLINE=86158754; PubMed=2420326;  
Kingston I.B., Anderson S.;  
"Sequences encoding two trypsin inhibitors occur in strikingly  
similar genomic environments.";  
Biochem. J. 233:443-450(1986).  
[4]  
SEQUENCE OF 34-97 FROM N.A.  
MEDLINE=84070725; PubMed=6580617;  
Anderson S., Kingston I.B.;  
"Isolation of a genomic clone for bovine pancreatic trypsin inhibitor  
by using a unique-sequence synthetic DNA probe.";

RL Proc. Natl. Acad. Sci. U.S.A. 80:6838-6842(1983).  
RN [5]  
RP SEQUENCE OF 36-93, AND DISULFIDE BONDS.  
RX MEDLINE=66083012; PubMed=5860161;  
RA Kassell B., Laskowski M.;  
RT "The basic trypsin inhibitor of bovine pancreas. V. The disulfide  
RT linkages.";  
RL Biochem. Biophys. Res. Commun. 20:463-468(1965).  
RN [6]  
RP SEQUENCE OF 36-93, AND DISULFIDE BONDS.  
RX MEDLINE=66171231; PubMed=5296424;  
RA Anderer F.A., Hornle S.;  
RT "The disulfide linkages in kallikrein inactivator of bovine lung.";  
RL J. Biol. Chem. 241:1568-1572(1966).  
RN [7]  
RP SEQUENCE OF 36-93, AND DISULFIDE BONDS.  
RX MEDLINE=68012003; PubMed=6053284;  
RA Chauvet J., Acher R.;  
RT "Covalent structure of a polypeptide inhibitor of trypsin (Kunitz and  
RT Northrop inhibitor).";  
RL Bull. Soc. Chim. Biol. 49:985-1000(1967).  
RN [8]  
RP SEQUENCE OF 36-93.  
RA Dlouha V., Pospisilova D., Meloun B., Sorm F.;  
RT "Sequence of residues 18-20 in pancreatic trypsin inhibitor.";  
RL Collect. Czech. Chem. Commun. 33:1363-1365(1968).  
RN [9]  
RP SEQUENCE OF 36-81.  
RC TISSUE=Adrenal chromaffin;  
RX MEDLINE=90211226; PubMed=2322242;  
RA Lewis R.V., Ray P., Coghill R., Kruggel W.;  
RT "Presence of pancreatic trypsin inhibitor in adrenal medullary  
RT chromaffin cells.";  
RL Biochem. Biophys. Res. Commun. 167:543-547(1990).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).  
RA Deisenhofer J., Steigemann W.;  
RT "Crystallographic refinement of the structure of bovine pancreatic  
RT trypsin inhibitor at 1.5-A resolution.";  
RL Acta Crystallogr. B 31:238-250(1975).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=70255230; PubMed=5447861;  
RA Huber R., Kukla D., Ruhlmann A., Epp O., Formanek H.;  
RT "The basic trypsin inhibitor of bovine pancreas. I. Structure  
RT analysis and conformation of the polypeptide chain.";  
RL Naturwissenschaften 57:389-392(1970).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF MUTANT GLY-70.  
RX MEDLINE=91332906; PubMed=1714504;  
RA Housset D., Kim K.-S., Fuchs J., Woodward C., Wlodawer A.;  
RT "Crystal structure of a Y35G mutant of bovine pancreatic trypsin  
RT inhibitor.";  
RL J. Mol. Biol. 220:757-770(1991).  
RN [13]  
RP STRUCTURE BY NMR.  
RX MEDLINE=93021116; PubMed=1383552;  
RA Berndt K.D., Guntert P., Orbons L.P.M., Wuethrich K.;  
RT "Determination of a high-quality nuclear magnetic resonance solution  
RT structure of the bovine pancreatic trypsin inhibitor and comparison  
RT with three crystal structures.";  
RL J. Mol. Biol. 227:757-775(1992).  
CC -|- FUNCTION: Inhibits trypsin, kallikrein, chymotrypsin, and plasmin.  
CC -|- SUBCELLULAR LOCATION: Secreted.  
CC -|- PHARMACEUTICAL: Available under the name Trasylol (Mile). Used  
CC for inhibiting coagulation so as to reduce blood loss during  
CC bypass surgery.  
CC -|- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
CC -|- DATABASE: NAME=Trasylol; NOTE=Clinical information on Trasylol;  
CC WWW="http://www.trasylol.com/".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; M20934; AAD13685.1; --  
DR EMBL; M20930; AAD13685.1; JOINED.  
DR EMBL; M20932; AAD13685.1; JOINED.  
DR EMBL; X03365; CAA27062.1; ALT\_SEQ.  
DR EMBL; X03365; CAA27063.1; --  
DR EMBL; X05274; CAA28886.1; --  
DR PIR; S00277; TIBO.  
DR PDB; 1TPA; 14-MAR-85.  
DR PDB; 2PTC; 14-MAR-85.  
DR PDB; 4PTI; 16-APR-87.  
DR PDB; 5PTI; 15-JAN-90.  
DR PDB; 6PTI; 16-JAN-88.  
DR PDB; 7PTI; 15-APR-91.  
DR PDB; 8PTI; 15-OCT-91.  
DR PDB; 9PTI; 15-OCT-92.  
DR PDB; 2TGP; 14-MAR-85.  
DR PDB; 2TPI; 15-JUL-91.  
DR PDB; 3TPI; 14-MAR-85.  
DR PDB; 4TPI; 08-NOV-85.  
DR PDB; 1AAL; 31-OCT-93.  
DR PDB; 1BPT; 15-JAN-93.  
DR PDB; 1BTI; 31-OCT-93.  
DR PDB; 1FAN; 31-OCT-93.  
DR PDB; 1NAG; 31-OCT-93.  
DR PDB; 2KAI; 15-OCT-91.  
DR PDB; 1PIT; 31-JAN-94.  
DR PDB; 1BRB; 31-JUL-94.  
DR PDB; 1BPI; 03-JUN-95.  
DR PDB; 1MTN; 17-AUG-96.  
DR PDB; 1BTH; 24-DEC-97.  
DR PDB; 1CBW; 23-JUL-97.  
DR PDB; 1B0C; 11-NOV-98.  
DR PDB; 1BZ5; 11-NOV-98.  
DR PDB; 3TGI; 23-DEC-98.  
DR PDB; 3TGJ; 23-DEC-98.  
DR PDB; 1FAK; 03-DEC-99.  
DR PDB; 1BHC; 16-SEP-98.  
DR PDB; 1BZX; 29-DEC-99.  
DR PDB; 1CO7; 07-JAN-03.  
DR PDB; 1D0D; 09-SEP-00.  
DR PDB; 1F5R; 04-JUL-01.  
DR PDB; 1F7Z; 04-JUL-01.  
DR PDB; 1FY8; 04-JUL-01.  
DR PDB; 1JV8; 27-FEB-02.  
DR PDB; 1JV9; 27-FEB-02.  
DR PDB; 1K6U; 19-DEC-01.  
DR PDB; 1LD5; 11-SEP-02.  
DR PDB; 1LD6; 11-SEP-02.  
DR PDB; 1QLQ; 03-FEB-00.  
DR PDB; 2HEX; 11-MAR-03.  
DR PDB; 3BTD; 13-MAR-00.  
DR PDB; 3BTE; 15-MAR-00.  
DR PDB; 3BTF; 13-MAR-00.  
DR PDB; 3BTG; 13-MAR-00.  
DR PDB; 3BTH; 13-MAR-00.  
DR PDB; 3BTK; 13-MAR-00.  
DR PDB; 3BTM; 13-MAR-00.  
DR PDB; 3BTQ; 13-MAR-00.  
DR PDB; 3BTT; 13-MAR-00.  
DR PDB; 3BTW; 13-MAR-00.  
DR PDB; 3TGK; 04-JUL-01.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR Pfam; PF00014; Kunitz\_BPTI; 1.  
DR PRINTS; PR00759; BASICTPASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 1.  
DR SMART; SM00131; KU; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
KW Serine protease inhibitor; Signal; Pharmaceutical; 3D-structure.  
FT SIGNAL 1 21 POTENTIAL.  
FT PROPEP 22 35  
FT CHAIN 36 93 PANCREATIC TRYPSIN INHIBITOR.  
FT PROPEP 94 100  
FT DOMAIN 40 90 BPTI/KUNITZ INHIBITOR.  
FT DISULFID 40 90  
FT DISULFID 49 73  
FT DISULFID 65 86  
FT ACT\_SITE 50 51 REACTIVE BOND (TRYPSIN).  
FT HELIX 38 41  
FT STRAND 53 59  
FT TURN 60 63  
FT STRAND 64 70

Query Match 76.6%; Score 36; DB 1; Length 100;  
Best Local Similarity 60.0%; Pred. No. 0.33;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10  
||| |||  
Db 47 GPCXARIIRY 56

## RESULT 6

BPT2 BOVIN

ID\_BPT2\_BOVIN STANDARD; PRT; 100 AA.

AC P04815;

DT 13-AUG-1987 (Rel. 05, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Spleen trypsin inhibitors I, II, and III precursor (SI).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87283904; PubMed=2441071;

RA Creighton T.E., Charles I.G.;

RT "Sequences of the genes and polypeptide precursors for two bovine

RT protease inhibitors.";

RL J. Mol. Biol. 194:11-22(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=88295740; PubMed=2456884;

RA Creighton T.E., Charles I.G.;

RT "Biosynthesis, processing, and evolution of bovine pancreatic trypsin

RT inhibitor.";

RL Cold Spring Harb. Symp. Quant. Biol. 52:511-519(1987).

RN [3]

RP SEQUENCE OF 34-97 FROM N.A. (SI-I).

RX MEDLINE=86158754; PubMed=2420326;

RA Kingston I.B., Anderson S.;

RT "Sequences encoding two trypsin inhibitors occur in strikingly

RT similar genomic environments.";

RL Biochem. J. 233:443-450(1986).

RN [4]

RP SEQUENCE OF 34-99 (SI-I).

RX MEDLINE=89076531; PubMed=2462435;

RA Fioretti E., Angeletti M., Fiorucci L., Barra D., Bossa F., Ascoli F.;

RT "Aprotinin-like isoinhibitors in bovine organs.";

RL Biol. Chem. Hoppe-Seyler 369:37-42(1988).

RN [5]

RP SEQUENCE OF 34-99 (SI-I).

RX MEDLINE=88007630; PubMed=3654647;

RA Barra D., Simmaco M., Bossa F., Fioretti E., Angeletti M., Ascoli F.;

RT "Primary structure of a protease isoinhibitor from bovine spleen. A

RT possible intermediate in the processing of the primary gene

RT product.";

RL J. Biol. Chem. 262:13916-13919(1987).

[6]  
SEQUENCE OF 36-93 (SI-II).  
MEDLINE=86008178; PubMed=2413011;  
Fioretti E., Iacopino G., Angeletti M., Barra D., Bossa F., Ascoli F.;  
"Primary structure and antiproteolytic activity of a Kunitz-type  
inhibitor from bovine spleen.";  
J. Biol. Chem. 260:11451-11455(1985).  
[7]  
SEQUENCE OF 36-97 (SI-III).  
MEDLINE=91098258; PubMed=1986787;  
Barra D., Fioretti E., Angeletti M., Maras B., Bossa F., Ascoli F.;  
"Proteinase isoinhibitors from bovine spleen: primary structure of an  
intermediate in the processing of the precursor.";  
Biochim. Biophys. Acta 1076:143-147(1991).  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M20935; AAA51418.1; -.  
EMBL; M20931; AAA51418.1; JOINED.  
EMBL; M20933; AAA51418.1; JOINED.  
EMBL; X05275; CAA28887.1; -.  
EMBL; X06685; CAA29881.1; -.  
EMBL; X03366; CAA27064.1; ALT\_SEQ.  
EMBL; X03366; CAA27065.1; -.  
PIR; S00274; TIBOSP.  
HSSP; P00974; 1BPI.  
InterPro; IPR002223; Kunitz\_BPTI.  
Pfam; PF00014; Kunitz\_BPTI; 1.  
PRINTS; PR00759; BASICPTASE.  
ProDom; PD000222; Kunitz\_BPTI; 1.  
SMART; SM00131; KU; 1.  
PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
PROSITE; PS0279; BPTI\_KUNITZ\_2; 1.  
Serine protease inhibitor; Signal.  
SIGNAL 1 21 POTENTIAL.  
PROPEP 22 33 SPLEEN TRYPSIN INHIBITOR I.  
CHAIN 34 99 SPLEEN TRYPSIN INHIBITOR II.  
CHAIN 36 93 SPLEEN TRYPSIN INHIBITOR III.  
CHAIN 36 97 BPTI/KUNITZ INHIBITOR.  
DOMAIN 40 90 BY SIMILARITY.  
DISULFID 40 90 BY SIMILARITY.  
DISULFID 49 73 BY SIMILARITY.  
DISULFID 65 86 BY SIMILARITY.  
ACT SITE 50 51 REACTIVE BOND (BY SIMILARITY).  
SEQUENCE 100 AA; 10843 MW; 39069734B8ACF4E3 CRC64;  
  
ery Match 76.6%; Score 36; DB 1; Length 100;  
st Local Similarity 60.0%; Pred. No. 0.33;  
tches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
1 GPCXXXFIRY 10  
||| |||  
47 GPCAKMIRY 56  
  
ILT 7  
BOMMO STANDARD; PRT; 55 AA.  
P81902;  
28-FEB-2003 (Rel. 41, Created)  
28-FEB-2003 (Rel. 41, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Trypsin inhibitor (Cocoon shell-associated trypsin inhibitor) (CSTI).  
Bombyx mori (Silk moth).  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE, AND CHARACTERIZATION.  
RC STRAIN=Asahi;  
RX MEDLINE=99115431; PubMed=9914483;  
RA Kurioka A., Yamazaki M., Hirano H.;  
RT "Primary structure and possible functions of a trypsin inhibitor of  
RT Bombyx mori.";  
RL Eur. J. Biochem. 259:120-126(1999).  
CC -!- FUNCTION: This cocoon shell-associated protein inhibits trypsin  
CC Activity by forming a low-dissociation complex with trypsin. May  
CC play an important part in regulating proteolytic activity in the  
CC silk gland or protecting silk proteins from degradation during  
CC histolysis.  
CC -!- SUBUNIT: Monomer.  
CC -!- TISSUE SPECIFICITY: Expressed exclusively in the middle silk  
CC gland.  
CC -!- DEVELOPMENTAL STAGE: Expression differentially regulated in the  
CC middle silk glands during the final stage of larval growth with  
CC highest expression before the onset of spinning.  
CC -!- MASS SPECTROMETRY: MW=6658; METHOD=Electrospray.  
CC -!- MISCELLANEOUS: Has an isoelectric point of 4.3.  
CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
DR HSSP; P31713; 1SHP.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR Pfam; PF00014; Kunitz\_BPTI; 1.  
DR PRINTS; PR00759; BASICPTASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 1.  
DR SMART; SM00131; KU; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
DR PROSITE; PS0279; BPTI\_KUNITZ\_2; 1.  
KW Developmental protein; Serine protease inhibitor.  
FT DISULFID 4 54 BY SIMILARITY.  
FT DISULFID 13 37 BY SIMILARITY.  
FT DISULFID 29 50 BY SIMILARITY.  
FT ACT\_SITE 14 15 REACTIVE BOND (TRYPSIN) (BY  
FT SIMILARITY).  
SQ SEQUENCE 55 AA; 6027 MW; C2739BB8F2BB6E59 CRC64;  
  
Query Match 74.5%; Score 35; DB 1; Length 55;  
Best Local Similarity 60.0%; Pred. No. 0.31;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 GPCXXXFIRY 10  
||| |||  
Db 11 GPCCKGSFPRY 20  
  
RESULT 8  
IBPI\_TACTR  
ID IBPI\_TACTR STANDARD; PRT; 61 AA.  
AC P16044;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Proteinase inhibitor (BPI-type).  
OS Tachypleus tridentatus (Japanese horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Tachypleus.  
OX NCBI\_TaxID=6853;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Hemocyte;  
RX MEDLINE=88007472; PubMed=3308864;  
RA Nakamura T., Hirai T., Tokunaga F., Kawabata S., Iwanaga S.;  
RT "Purification and amino acid sequence of Kunitz-type protease  
RT inhibitor found in the hemocytes of horseshoe crab (Tachypleus  
RT tridentatus).";  
RL J. Biochem. 101:1297-1306(1987).  
CC -!- FUNCTION: Inhibitor of trypsin and chymotrypsin.  
CC -!- SUBCELLULAR LOCATION: Secreted.



CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
DR PIR; A26923; TIFCBP.  
DR HSSP; P00974; IBPI.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR Pfam; PF00014; Kunitz\_BPTI; 1.  
DR PRINTS; PR00759; BASICPTASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 1.  
DR SMART; SM00131; KU; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
KW Serine protease inhibitor.

FT DISULFID 8 58 BY SIMILARITY.  
FT DISULFID 17 41 BY SIMILARITY.  
FT DISULFID 33 54 BY SIMILARITY.  
FT ACT SITE 18 19 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 61 AA; 6825 MW; 730E82CDD0653E48 CRC64;

Query Match 74.5%; Score 35; DB 1; Length 61;  
Best Local Similarity 60.0%; Pred. No. 0.34;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |  
Db 15 GPCRAGPKRY 24

## RESULT 9

IP52\_ANESU STANDARD; PRT; 62 AA.  
AC P10280;

DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protease inhibitor 5 II (SA5 II).

OS Anemonia sulcata (Snake-locks sea anemone).

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;

OC Nynanthaeae; Actiniidae; Anemonia.

OX NCBI\_TaxID=6108;

[1]

RP SEQUENCE OF 1-59.

RA Wunderer G., Machleidt W., Fritz H.;

RT "The broad-specificity proteinase inhibitor 5 II from the sea anemone

RT Anemonia sulcata.";

RL Meth. Enzymol. 80:816-820 (1981).

RN [2]

RP SEQUENCE.

RA Krebs H.C., Habermehl G.G.;

RT "Isolation and structural determination of a hemolytic active peptide

RT from the sea anemone Metridium senile.";

RL Naturwissenschaften 74:395-396 (1987).

CC -!- FUNCTION: Inhibitor of kallikreins.

CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

DR PIR; S07451; S07451.

DR HSSP; P31713; 15HP.

DR InterPro; IPR002223; Kunitz\_BPTI.

DR Pfam; PF00014; Kunitz\_BPTI; 1.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz\_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.

KW Serine protease inhibitor.

FT DISULFID 5 55 BY SIMILARITY.

FT DISULFID 14 38 BY SIMILARITY.

FT DISULFID 30 51 BY SIMILARITY.

FT ACT SITE 15 16 REACTIVE BOND (BY SIMILARITY).

FT VARIANT 13 13 P -> R.

FT VARIANT 16 16 A -> G.

FT VARIANT 17 17 R -> G.

FT VARIANT 25 25 S -> L.

FT VARIANT 28 28 K -> R.

FT VARIANT 39 39 G -> R.

SQ SEQUENCE 62 AA; 6937 MW; 7262D028CA567BC8 CRC64;

Query Match 74.5%; Score 35; DB 1; Length 62;  
Best Local Similarity 60.0%; Pred. No. 0.35;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |  
Db 12 GPCRARFPY 21

## RESULT 10

IBPC\_BOVIN

ID IBPC\_BOVIN STANDARD; PRT; 67 AA.

AC P00976;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Colostrum trypsin inhibitor (Colostrum BPI).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE.

RA Cechova D., Jonakova V., Sorm F.;

RT "Primary structure of trypsin inhibitor from cow colostrum (component

RT B2).";

RL Collect. Czech. Chem. Commun. 36:3342-3357 (1971).

RN [2]

RP DISULFIDE BONDS.

RA Cechova D., Ber E.;

RT "Disulfide bonds of trypsin inhibitor from cow colostrum.";

RL Collect. Czech. Chem. Commun. 39:680-688 (1974).

RN [3]

RP CHARACTERIZATION.

RA Cechova D., Muszynska G.;

RT "Role of lysine 18 in active center of cow colostrum trypsin

RT inhibitor.";

RL FEBS Lett. 8:84-86 (1970).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

DR PIR; A01207; TIBOC.

DR HSSP; P02760; 1BIK.

DR InterPro; IPR002223; Kunitz\_BPTI.

DR Pfam; PF00014; Kunitz\_BPTI; 1.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz\_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.

KW Serine protease inhibitor; Glycoprotein.

FT DISULFID 8 58

FT DISULFID 17 41

FT DISULFID 33 54

FT CARBOHYD 27 27

FT ACT SITE 18 19

SQ SEQUENCE 67 AA; 7511 MW; E2B2093B7CD207CD CRC64;

Query Match 74.5%; Score 35; DB 1; Length 67;  
Best Local Similarity 50.0%; Pred. No. 0.38;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |  
Db 15 GPCXAALLRY 24

## RESULT 11

MCPI\_MELCP

ID MCPI\_MELCP STANDARD; PRT; 197 AA.

AC P82968;

DT 28-FEB-2003 (Rel. 41, Created)



28-FEB-2003 (Rel. 41, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Protease inhibitor (McpI).  
Melithaea caledonica.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Gorgonacea;  
Scleraxonia; Melithaeidae; Melithaea.  
NCBI\_TaxID=156534;  
[1]  
SEQUENCE, AND VARIANT VAL-195.  
Peduzzi J., Longeon A., Guyot M., Barthelemy M.;  
"Amino acid sequence and kinetic properties of a four-domain proteases  
inhibitor from the gorgonian Melithaea caledonica.";  
Eur. J. Biochem. 0:0-0(2001).  
-!- FUNCTION: INHIBITS TRYPSIN, KALLIKREIN, SUBTILISIN CARLSBERG,  
HUMAN LEUKOCYTE ELASTASE, PORCINE PANCREATIC ELASTASE AND  
CHYMOTRYPSIN. TWO DOMAINS ARE FOR THE INHIBITION OF CHYMOTRYPSIN.  
-!- SIMILARITY: Contains 3 Kazal-like domains.  
-!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
HSSP; P31713; 1SHP.  
InterPro; IPR002350; kazal.  
InterPro; IPR002223; Kunitz\_BPTI.  
Pfam; PF00050; kazal; 3.  
PRINTS; PF00014; Kunitz\_BPTI; 1.  
ProDom; PD000222; Kunitz\_BPTI; 1.  
SMART; SM00131; KU; 1.  
PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
PROSITE; PS0279; BPTI\_KUNITZ\_2; 1.  
PROSITE; PS00282; KAZAL; FALSE\_NEG.  
Serine protease inhibitor; Repeat.  
DOMAIN 1 45 KAZAL-LIKE 1.  
DOMAIN 49 93 KAZAL-LIKE 2.  
DOMAIN 97 141 KAZAL-LIKE 3.  
DOMAIN 144 194 BPTI/KUNITZ INHIBITOR.  
ACT\_SITE 7 8 REACTIVE BOND 1 (BY SIMILARITY).  
ACT\_SITE 55 56 REACTIVE BOND 2 (BY SIMILARITY).  
ACT\_SITE 103 104 REACTIVE BOND 3 (BY SIMILARITY).  
ACT\_SITE 154 155 REACTIVE BOND 4 (BY SIMILARITY).  
DISULFID 1 31 BY SIMILARITY.  
DISULFID 5 24 BY SIMILARITY.  
DISULFID 13 45 BY SIMILARITY.  
DISULFID 49 79 BY SIMILARITY.  
DISULFID 53 72 BY SIMILARITY.  
DISULFID 61 93 BY SIMILARITY.  
DISULFID 97 127 BY SIMILARITY.  
DISULFID 101 120 BY SIMILARITY.  
DISULFID 109 141 BY SIMILARITY.  
DISULFID 144 194 BY SIMILARITY.  
DISULFID 153 177 BY SIMILARITY.  
DISULFID 169 190 BY SIMILARITY.  
VARIANT 195 195 M -> V.  
SEQUENCE 197 AA; 21248 MW; 199D08A489879579 CRC64;  
ery Match 74.5%; Score 35; DB 1; Length 197;  
st Local Similarity 60.0%; Pred. No. 1.1;  
tches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
1 GPCXXXFIRY 10  
151 GPCRGAFRRY 160  
HLT 12  
SALTY STANDARD; PRT; 330 AA.  
P26395;  
01-AUG-1992 (Rel. 23, Created)  
01-AUG-1992 (Rel. 23, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Rfbi protein.  
RFBI OR STM2093.  
Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2;  
RX MEDLINE=91260454; PubMed=1710759;  
RA Jiang X.-M., Neal B., Santiago F., Lee S.J., Romana L.K., Reeves P.R.;  
RT "Structure and sequence of the rfb (O antigen) gene cluster of  
RT Salmonella serovar typhimurium (strain LT2).";  
RL Mol. Microbiol. 5:695-713(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";  
RL Nature 413:852-856(2001).  
CC -!- PATHWAY: Lipopolysaccharide O antigen biosynthesis.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X56793; CAA40119.1; -.  
DR EMBL; AE008792; AAL20997.1; -.  
DR PIR; S15303; S15303.  
DR HSSP; P00235; 1FRR.  
DR StyGene; SG10451; rfbI.  
DR InterPro; IPR006058; 2Fe2S fd BS.  
DR InterPro; IPR008333; FAD\_binding\_6.  
DR InterPro; IPR001041; Ferredoxin.  
DR InterPro; IPR001709; FPN cyt redctse.  
DR InterPro; IPR001433; Oxred\_FAD/NAD(P).  
DR InterPro; IPR001221; Phe\_hydroxylase.  
DR Pfam; PF00970; FAD\_binding\_6; 1.  
DR Pfam; PF00111; fer2; 1.  
DR Pfam; PF00175; NAD\_binding\_1; 1.  
DR PRINTS; PR00371; FPNCR.  
DR PRINTS; PR00410; PHEHYDRXLASE.  
DR PROSITE; PS00197; 2FE2S-FERREDOXIN; UNKNOWN\_1.  
KW Lipopolysaccharide biosynthesis; Complete proteome.  
SQ SEQUENCE 330 AA; 36582 MW; EFC1BEC17A0CC82D CRC64;  
Query Match 74.5%; Score 35; DB 1; Length 330;  
Best Local Similarity 66.7%; Pred. No. 1.8;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GPCXXXFIR 9  
Db 184 GPCGTFIR 192  
RESULT 13  
IATR\_SHEEP  
ID IATR\_SHEEP STANDARD; PRT; 123 AA.  
AC P13371;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Inter-alpha-trypsin inhibitor (ITI) (GIK-14) (Inhibitory fragment of  
DE ITI) (Fragment).  
OS Ovis aries (Sheep), and  
OS Capra hircus (Goat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940, 9925;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=Sheep;  
RX MEDLINE=87299012; PubMed=2441725;  
RA Rasp G., Hochstrasser K., Wachter E., Reisinger P.W.M.;  
RT "The amino-acid sequence of the trypsin-released inhibitor from sheep  
inter-alpha-trypsin inhibitor.";  
RL Biol. Chem. Hoppe-Seyler 368:727-731(1987).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=C.hircus;  
RX MEDLINE=90105540; PubMed=2481505;  
RA Rasp G., Hochstrasser K., Gerl C., Wachter E.;  
RT "Primary structure of a proteinase inhibitor released from goat serum  
inter-alpha-trypsin inhibitor.";  
RL Biochim. Biophys. Acta 999:335-337(1989).  
CC -!- FUNCTION: This inhibitory fragment, released from native ITI after  
limited proteolysis with trypsin, contains two homologous domains.  
CC Whereas the second domain is a strong inhibitor of trypsin, the  
CC first domain interacts weakly with PMN-granulocytic elastase and  
CC not at all with pancreatic elastase.  
CC -!- MISCELLANEOUS: The amino acid at position p2' (17) appears to  
CC determine the specificity of the inhibition of domain I.  
CC Inhibitors with methionine in this position interact weakly with  
CC chymotrypsin and elastase; those with leucine interact strongly.  
CC -!- SIMILARITY: Contains 2 BPTI/Kunitz inhibitor domains.  
DR PIR; A29652; A29652.  
DR HSSP; P02760; 1BIK.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR Pfam; PF00014; Kunitz\_BPTI\_2.  
DR PRINTS; PR00759; BASICTPASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 2.  
DR SMART; SM00131; KU; 2.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 2.  
DR PROSITE; PS02079; BPTI\_KUNITZ\_2; 2.  
KW Plasma; Glycoprotein; Serine protease inhibitor; Repeat.  
FT NON TER 1 1 BPTI/KUNITZ INHIBITOR 1.  
FT DOMAIN 5 55 BPTI/KUNITZ INHIBITOR-2.  
FT DISULFID 61 111 BY SIMILARITY.  
FT DISULFID 5 55 BY SIMILARITY.  
FT DISULFID 14 38 BY SIMILARITY.  
FT DISULFID 30 51 BY SIMILARITY.  
FT DISULFID 61 111 BY SIMILARITY.  
FT DISULFID 70 94 BY SIMILARITY.  
FT DISULFID 86 107 BY SIMILARITY.  
FT SITE 15 16 INHIBITORY (P1) (CHYMOTRYPSIN, ELASTASE).  
FT SITE 71 72 INHIBITORY (P1) (TRYPSIN).  
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .).  
FT NON TER 123 123  
SQ SEQUENCE 123 AA; 13686 MW; 295038173F22D2D1 CRC64;  
Query Match 72.3%; Score 34; DB 1; Length 123;  
Best Local Similarity 60.0%; Pred. No. 1.1;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GPCXXXFYRY 10  
Db 12 GPCGLGMFKRY 21  
RESULT 14  
TFP2 HUMAN STANDARD; PRT; 235 AA.  
AC P48307;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tissue factor pathway inhibitor 2 precursor (TFPI-2) (Placental  
protein 5) (pp5).

GN  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Placenta;  
RX MEDLINE=95204397; PubMed=7896752;  
RA Miyagi Y., Koshikawa N., Yasumitsu H., Miyagi E., Hirahara F.,  
Aoki I., Misugi K., Umeda M., Miyazaki K.;  
RT "cDNA cloning and mRNA expression of a serine proteinase inhibitor  
secreted by cancer cells: identification as placental protein 5 and  
tissue factor pathway inhibitor-2.";  
RL J. Biochem. 116:939-942(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=94211862; PubMed=8159751;  
RA Sprecher C.A., Kisiel W., Mathewes S., Foster D.C.;  
RT "Molecular cloning, expression, and partial characterization of a  
second human tissue-factor-pathway inhibitor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:3353-3357(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Maggi L.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21240334; PubMed=11342222;  
RA Kamei S., Kazama Y., Kuijper J.L., Foster D.C., Kisiel W.;  
RT "Genomic structure and promoter activity of the human tissue factor  
pathway inhibitor-2 gene.";  
RL Biochim. Biophys. Acta 1517:430-435(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Touchman J.W., Green E.D., Dickson M.C.,  
Blakesley R.W., Grimwood J., Schmutz J., Myers R.M.,  
Rodriguez A.C., Krzywinski M.I., Skalska U., Bouffard G.G.,  
Butterfield Y.S.N., Jones S.J.M., Marra M.A.;  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP PARTIAL SEQUENCE OF 23-35; 47-53 AND 133-146.  
RC TISSUE=Placenta;  
RX MEDLINE=88106628; PubMed=3276312;  
RA Buetzow R., Huhtala M.-L., Bohn H., Virtanen I., Seppaelae M.;  
RT "Purification and characterization of placental protein 5.";  
RL Biochem. Biophys. Res. Commun. 150:483-490(1988).  
RN [7]  
RP ERRATUM.  
RA Buetzow R., Huhtala M.-L., Bohn H., Virtanen I., Seppaelae M.;  
RL Biochem. Biophys. Res. Commun. 151:630-631(1988).  
CC -!- FUNCTION: Seems to inhibit trypsin, factor VIIa/tissue factor,  
CC weakly factor Xa. Has no effect on thrombin.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Umbilical vein endothelial cells, liver,





CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.  
CC -!- PTM: The precursor is proteolytically processed into two  
CC separately functioning proteins.  
CC -!- PTM: Alpha-1-microglobulin contains covalently linked brown-  
CC yellow chromophores (By similarity).  
CC -!- SIMILARITY: In the N-terminal section; belongs to the lipocalin  
CC family.  
CC -!- SIMILARITY: Contains 2 BPTI/Kunitz inhibitor domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U35642; AAB07599.1; -.  
CC PIR; S68149; TIBOB1.  
CC HSSP; P02760; 1BIK.  
CC InterPro; IPR002223; Kunitz\_BPTI.  
CC InterPro; IPR002345; Lipocalin.  
CC InterPro; IPR000566; Lipocalin\_cyFABP.  
CC Pfam; PF00014; Kunitz\_BPTI; 2.  
CC Pfam; PF00061; lipocalin; 1.  
CC PRINTS; PR00759; BASICPTASE.  
CC PRINTS; PR00179; LIPOCALIN.  
CC ProDom; PD000222; Kunitz\_BPTI; 2.  
CC SMART; SM00131; KU; 2.  
CC PROSITE; PS00280; BPTI\_KUNITZ\_1; 2.  
CC PROSITE; PS00279; BPTI\_KUNITZ\_2; 2.  
CC PROSITE; PS00213; LIPOCALIN; 1.  
KW Glycoprotein; Plasma; Signal; Serine protease inhibitor; Repeat;  
KW Lipocalin.  
FT SIGNAL 1 19 BY SIMILARITY.  
FT CHAIN 20 203 ALPHA-1 MICROGLOBULIN.  
FT CHAIN 206 352 INTER-ALPHA-TRYPSIN INHIBITOR LIGHT  
FT CHAIN CHAIN.  
FT DOMAIN 231 281 BPTI/KUNITZ INHIBITOR 1.  
FT DOMAIN 287 337 BPTI/KUNITZ INHIBITOR 2.  
FT BINDING 53 53 CHROMOPHORE (BY SIMILARITY).  
FT BINDING 111 111 CHROMOPHORE (BY SIMILARITY).  
FT BINDING 137 137 CHROMOPHORE (BY SIMILARITY).  
FT BINDING 149 149 CHROMOPHORE (BY SIMILARITY).  
FT DISULFID 91 188 BY SIMILARITY.  
FT DISULFID 231 281 BY SIMILARITY.  
FT DISULFID 240 264 BY SIMILARITY.  
FT DISULFID 256 277 BY SIMILARITY.  
FT DISULFID 287 337 BY SIMILARITY.  
FT DISULFID 296 320 BY SIMILARITY.  
FT DISULFID 312 333 BY SIMILARITY.  
FT SITE 241 242 INHIBITORY (P1) (CHYMOTRYPSIN, ELASTASE).  
FT SITE 297 298 INHIBITORY (P1) (TRYPSIN).  
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .).  
FT CONFLICT 209 209 T -> G (IN REF. 4).  
FT CONFLICT 217 217 A -> D (IN REF. 4).  
FT CONFLICT 268 268 G -> L (IN REF. 2 AND 3).  
FT CONFLICT 274 274 E -> Q (IN REF. 2 AND 3).  
FT CONFLICT 298 299 SY -> AF (IN REF. 2 AND 3).  
FT CONFLICT 330 330 E -> Q (IN REF. 2 AND 3).  
FT CONFLICT 346 346 E -> R (IN REF. 2 AND 3).  
SQ SEQUENCE 352 AA; 39235 MW; ED31C5CA02E70B19 CRC64;

Search completed: August 11, 2004, 13:08:06  
Job time : 8.61905 secs

Query Match 72.3%; Score 34; DB 1; Length 352;  
Best Local Similarity 60.0%; Pred. No. 3.2;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10  
||| |  
Db 238 GPCGLGPKRY 247



protein - protein search, using sw model

On: August 11, 2004, 13:04:18 ; Search time 31.9048 Seconds  
(without alignments)  
98.894 Million cell updates/sec

US-10-661-207-1

ject score 47

ience: 1 GPCXXXFIRY 10

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 1017041 seqs, 315518202 residues

l number of hits satisfying chosen parameters: 1017041

um DB seq length: 0

um DB seq length: 2000000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

base :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

lt	Score	Query	Match	Length	DB	ID	Description
1	37	78.7	133	6	Q8HZ45	Q8hz45	papio papio
2	36	76.6	76	5	Q968S8	Q968s8	galleria me
3	36	76.6	461	5	Q95S22	Q95s22	drosophila
4	36	76.6	763	5	Q9XZD0	Q9xzdo	drosophila
5	35	74.5	76	5	Q8T7L9	Q8t7l9	bombyx mori
6	35	74.5	324	16	Q8Z5I3	Q8z5i3	salmonella
7	35	74.5	580	5	Q193I5	Q193i5	caenorhabdi
8	35	74.5	2225	5	O45881	O45881	caenorhabdi
9	34	72.3	91	7	O19474	O19474	mus musculu
10	34	72.3	224	4	Q8NAK6	Q8nak6	homo sapien
11	34	72.3	253	11	Q62323	Q62323	mus musculu
12	34	72.3	272	7	Q860A4	Q860a4	mus caroli
13	34	72.3	272	7	Q860A1	Q860a1	mus abbotti
14	34	72.3	272	7	Q85ZZ9	Q85zz9	mus caroli
15	34	72.3	272	7	Q85ZZ6	Q85zz6	mus dunni (
16	34	72.3	273	7	Q8HWI2	Q8hw12	rattus norv

RESULT 1

Q8HZ45 ID Q8HZ45 PRELIMINARY; PRT; 133 AA.  
AC Q8HZ45;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Epididymal protease inhibitor 1.  
GN EPPIN.  
OS Papio papio (Guinea baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Papio.  
OX NCBI\_TaxID=100937;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Sivashanmugam P., O'Rand M.G., Richardson R.T.;  
RT "Characterization Of Mouse Eppin and a Gene Cluster of Similar  
RT Protease Inhibitors on Mouse Chromosome 2 Gene."  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY141973; AAN08507.1; -;  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR InterPro; IPR008197; WAP.  
DR Pfam; PF00014; Kunitz\_BPTI; 1.  
DR Pfam; PF00095; wap; 1.  
DR PRINTS; PR00759; BASICPTASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 1.  
DR SMART; SM00131; KU; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
KW Protease.  
SQ SEQUENCE 133 AA; 15277 MW; B33AEE57ECBEBE84 CRC64;

Query Match 78.7%; Score 37; DB 6; Length 133;  
Best Local Similarity 60.0%; Pred. No. 0.79;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

17 34 72.3 273 7 Q860A2 Q860a2 mus abbotti  
18 34 72.3 273 7 Q860A0 Q860a0 mus abbotti  
19 34 72.3 273 7 Q85ZZ8 Q85zz8 mus cookii  
20 34 72.3 273 7 Q85ZZ7 Q85zz7 mus cookii  
21 34 72.3 273 7 Q85ZZ5 Q85zz5 mus dunni  
22 34 72.3 273 7 Q85ZZ4 Q85zz4 mus musculu  
23 34 72.3 273 7 Q85ZZ2 Q85zz2 mus musculu  
24 34 72.3 273 7 Q85ZZ0 Q85zz0 mus musculu  
25 34 72.3 273 7 Q85ZY9 Q85zy9 mus musculu  
26 34 72.3 273 7 Q85ZY8 Q85zy8 mus musculu  
27 34 72.3 273 7 Q85ZY7 Q85zy7 mus pahari  
28 34 72.3 273 7 Q85ZY6 Q85zy6 mus platyth  
29 34 72.3 273 7 Q85ZY5 Q85zy5 mus platyth  
30 34 72.3 273 7 Q85ZY4 Q85zy4 mus platyth  
31 34 72.3 273 7 Q85ZY2 Q85zy2 mus setulos  
32 34 72.3 273 7 Q85ZY1 Q85zy1 mus setulos  
33 34 72.3 273 7 Q85ZY0 Q85zy0 mus shortri  
34 34 72.3 273 7 Q85ZX9 Q85zx9 mus shortri  
35 34 72.3 352 11 Q70160 Q70160 cavia porce  
36 34 72.3 384 7 Q95551 Q95551 mus musculu  
37 34 72.3 384 7 Q31202 Q31202 mus musculu  
38 34 72.3 384 7 Q8HNB4 Q8hnb4 mus musculu  
39 34 72.3 384 11 Q60678 Q60678 mus musculu  
40 34 72.3 395 11 Q8CGD8 Q8cgd8 mus musculu  
41 34 72.3 410 10 Q9C6V1 Q9c6v1 arabidopsis  
42 34 72.3 431 10 Q8RWY1 Q8rwy1 arabidopsis  
43 33 70.2 82 5 Q8MVB4 Q8mbv4 ixodes scap  
44 33 70.2 407 16 Q8A456 Q8a456 bacteroides  
45 33 70.2 605 5 Q965Q9 Q965q9 caenorhabdi



[6]  
SEQUENCE FROM N.A.  
FlyBase;  
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; AY060990; AAL28538.1; -  
EMBL; AE003804; AAM68494.1; -  
FlyBase; FBgn0026721; fat-spondin.  
GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
InterPro; IPR002223; Kunitz\_BPTI.  
InterPro; IPR000884; TSPI.  
Pfam; PF00014; Kunitz\_BPTI; 1.  
Pfam; PF00090; tsp\_1; 4.  
PRINTS; PR00759; BASICTPASE.  
ProDom; PD000222; Kunitz\_BPTI; 1.  
PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
PROSITE; PS50092; TSPI; 4.  
Protease inhibitor; Serine protease inhibitor.  
SEQUENCE 461 AA; 51517 MW; B77D07F41DD681B1 CRC64;  
  
Query Match 76.6%; Score 36; DB 5; Length 461;  
Best Local Similarity 50.0%; Pred. No. 4.3;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
1 GPCXXXFIRY 10  
||| :|||  
349 GPCRGTMYRY 358  
  
JLT 4  
Q9XZD0 PRELIMINARY; PRT; 763 AA.  
Q9XZD0;  
01-NOV-1999 (TrEMBLrel. 12, Created)  
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
Fat-spondin protein.  
FAT-SPONDIN OR CG6953.  
Drosophila melanogaster (Fruit fly).  
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=Berkley;  
MEDLINE=20196006; PubMed=10731132;  
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OREGON-R;  
RA DaRocha S., Baumgartner S.;  
RT "Fat-Spondin, a Drosophila member of the Spondin family, is highly  
RT expressed in fat body and hemocytes.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003804; AAF57910.1; -  
DR EMBL; AF135119; AAD31715.1; -  
DR HSSP; P00974; IRTI.  
DR FlyBase; FBgn0026721; fat-spondin.  
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR InterPro; IPR002861; Reeler.  
DR InterPro; IPR000884; TSPI.  
DR Pfam; PF00014; Kunitz\_BPTI; 1.  
DR Pfam; PF02014; Reeler; 1.  
DR Pfam; PF00090; tsp\_1; 4.  
DR PRINTS; PR00759; BASICTPASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 1.  
DR SMART; SM00131; KU; 1.  
DR SMART; SM00209; TSPI; 4.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
DR PROSITE; PS50092; TSPI; 4.  
KW Protease inhibitor; Serine protease inhibitor.  
SQ SEQUENCE 763 AA; 84945 MW; 3292DEDD2CFE4DAB CRC64;  
  
Query Match 76.6%; Score 36; DB 5; Length 763;  
Best Local Similarity 50.0%; Pred. NO. 7;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 GPCXXXFIRY 10  
||| :|||  
Db 651 GPCRGTMYRY 660  
  
RESULT 5  
Q8T7L9  
ID Q8T7L9 PRELIMINARY; PRT; 76 AA.  
AC Q8T7L9;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Kazal-type serine proteinase inhibitor 1.  
GN SPII.  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21873253; PubMed=11881808;  
RA Nirmala X., Mita K., Vanisree V., Zurovec M., Sehna F.;  
RT "Identification of four small molecular mass proteins in the silk of  
RT Bombyx mori.";  
RL Insect Mol. Biol. 10:437-445(2001).  
DR EMBL; AF352583; AAL83944.1; -  
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
DR InterPro; IPR002223; Kunitz\_BPTI.



```
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS02279; BPTI_KUNITZ_2; 1.
KW Protease inhibitor; Serine protease inhibitor.
SQ SEQUENCE 76 AA; 8445 MW; AA3F97D373535A3B CRC64;

Query Match      74.5%; Score 35; DB 5; Length 76;
Best Local Similarity 60.0%; Pred. No. 1.4;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10
   |||  |||
Db 32 GPCXGSPRY 41

RESULT 6
Q825I3 PRELIMINARY; PRT; 324 AA.
AC Q825I3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative reductase RfbI.
GN RFB1 OR STY2303 OR T0779.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627273; CAD02456.1; -.
DR EMBL; AE016836; AA068470.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006058; 2Fe2S fd BS.
DR InterPro; IPR008333; FAD binding_6.
DR InterPro; IPR01041; Ferredoxin.
DR InterPro; IPR01709; FPN cyt redctse.
DR InterPro; IPR01433; Oxred_FAD/NAD(P).
DR InterPro; IPR01221; Phe hydroxylase.
DR Pfam; PF00970; FAD binding_6; 1.
DR Pfam; PF01111; fer2; 1.
DR Pfam; PF00175; NAD binding_1; 1.
DR PRINTS; PR00371; PFNCR.
DR PRINTS; PR00410; PHEHYDRXLASE.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW Hypothetical protein; Complete proteome.
```

```
SQ SEQUENCE 324 AA; 35884 MW; E3A556150F5CE199 CRC64;

Query Match      74.5%; Score 35; DB 16; Length 324;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 9
   |||  |||
Db 184 GPCGTFYR 192

RESULT 7
Q19315 PRELIMINARY; PRT; 580 AA.
AC Q19315;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F10F2.8 protein.
GN F10F2.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Coles L.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z35598; CAA84654.1; -.
DR PIR; T20716; T20716.
DR WormPep; F10F2.8; CE00953.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM0034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 580 AA; 66437 MW; 08687AC6C55A7FDF CRC64;

Query Match      74.5%; Score 35; DB 5; Length 580;
Best Local Similarity 50.0%; Pred. No. 9.2;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GPCXXXFYRY 10
   |||  |||
Db 395 GPCPAGFLQY 404

RESULT 8
Q45881 PRELIMINARY; PRT; 2225 AA.
AC Q45881;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE W01F3.3 protein.
GN W01F3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Cummings P.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```



```
MEDLINE=99069613; PubMed=9851916;
none;
"Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
Science 282:2012-2018(1998).
EMBL; Z92815; CAB07294.1; -.
PIR; T26063; T26063.
HSSP; P31713; 1SHP.
WormPep; W01F3.3; CE16531.
GO; GO:0004867; F:serine protease inhibitor activity; IEA.
InterPro; IPR002223; Kunitz_BPTI.
InterPro; IPR000716; Thyroglobulin_1.
InterPro; IPR006150; Worm_repeat_1.
Pfam; PF00014; Kunitz_BPTI; 10.
PRINTS; PR00759; BASICPTASE.
ProDom; PD000222; Kunitz_BPTI; 10.
SMART; SM00131; KU; 10.
SMART; SM00211; TY; 1.
SMART; SM00289; WR1; 4.
PROSITE; PS00280; BPTI_KUNITZ_1; 8.
PROSITE; PS50279; BPTI_KUNITZ_2; 10.
PROSITE; PS00484; THYROGLOBULIN_1; 1.
Protease inhibitor; Serine protease inhibitor.
SEQUENCE 2225 AA; 242198 MW; A5DD8AE9D2A7B02A CRC64;

Query Match 74.5%; Score 35; DB 5; Length 2225;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GPCXXXFIRY 10
|||||
470 GPCGHSFQRY 479

JLT 9
174
O19474 PRELIMINARY; PRT; 91 AA.
O19474;
01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
MHC class I c5/gil (Tla) (Fragment).
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE=85206117; PubMed=3997208;
Rogers J.H.;
"Family organization of mouse H-2 class I genes.";
Immunogenetics 21:343-353(1985).
EMBL; M14828; AAA39688.1; -.
PIR; I68704; I68704.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0006955; P:immune response; IEA.
InterPro; IPR001039; MHC_I.
Pfam; PF00129; MHC_I; 1.
PRINTS; PR01638; MHCCLASSI.
ProDom; PD000050; MHC_I; 1.
NON TER 1.
SEQUENCE 91 AA; 10671 MW; 4A5571396359B05F CRC64;

Query Match 72.3%; Score 34; DB 7; Length 91;
Best Local Similarity 50.0%; Pred. No. 2.8;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

1 GPCXXXFIRY 10
|||||
71 GPCKDSLRLY 80
```

```
RESULT 10
Q8NAK6
ID Q8NAK6 PRELIMINARY; PRT; 224 AA.
AC Q8NAK6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ35180.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK092499; BAC03906.1; -.
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR008296; TFPI.
DR Pfam; PF00014; Kunitz_BPTI; 3.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 3.
DR SMART; SM00131; KU; 3.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 2.
DR PROSITE; PS50279; BPTI_KUNITZ_2; 3.
DR PIRSF; PIRSF001620; TFPI; 1.
KW Hypothetical protein; Protease inhibitor; Serine protease inhibitor.
SQ SEQUENCE 224 AA; 25795 MW; F586328C31344765 CRC64;

Query Match 72.3%; Score 34; DB 4; Length 224;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPCXXXFIRY 10
|||||
DB 32 GPCRALLRY 41

RESULT 11
Q62323
ID Q62323 PRELIMINARY; PRT; 253 AA.
AC Q62323;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TL antigen (Fragment).
GN H2-T18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86010114; PubMed=3840195;
RA Chen Y.T., Obata Y., Stockert E., Old L.J.;
RT "Thymus-leukemia (TL) antigens of the mouse.";
RL J. Exp. Med. 162:1134-1148(1985).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; X03052; CAA26860.1; -.
DR PIR; I48850; I48850.
DR HSSP; P01901; 1BQH.
DR MGD; MGI:95950; H2-T18.
```

DR GO:0016021; C:integral to membrane; IEA.  
DR GO:0006955; P:immune response; IEA.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig\_c1.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00129; MHC\_I; 1.  
DR PRINTS; PR01638; MHCCLASSI.  
DR ProDom; PD000050; MHC\_I; 1.  
DR SMART; SM00407; IGc1; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Glycoprotein; Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 253 AA; 28532 MW; 5D0012815EC77FB1 CRC64;

Query Match 72.3%; Score 34; DB 11; Length 253;  
Best Local Similarity 50.0%; Pred. No. 7.3;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| :||  
Db 57 GPCCKOSLLRY 66

## RESULT 12

ID Q860A4 PRELIMINARY; PRT; 272 AA.  
AC Q860A4;

DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE MHC class I antigen (Fragment).  
GN MUCATL.

OS Mus caroli (Wild mouse) (Ricefield mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10089;  
RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Tail;

RX MEDLINE=22359207; PubMed=12471122;

RA Davis B.K., Cook R.G., Rich R.R., Rodgers J.R.;

RT "Hyperconservation of the putative antigen recognition site of the MHC  
class I-b molecule TL in the subfamily Murinae: evidence that thymus  
leukemia antigen is an ancient mammalian gene.";

RL J. Immunol. 169:6890-6899(2002).

DR EMBL; AY144137; AAP23911.1; -.

DR EMBL; AY144136; AAP23911.1; JOINED.

DR GO:0016020; C:membrane; IEA.

DR GO:0006955; P:immune response; IEA.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig\_c1.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR001039; MHC\_I.

DR Pfam; PF00047; Ig; 1.

DR Pfam; PF00129; MHC\_I; 1.

DR PRINTS; PR01638; MHCCLASSI.

DR ProDom; PD000050; MHC\_I; 1.

DR SMART; SM00407; IGc1; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.

DR PROSITE; PS00290; IG\_MHC; 1.

FT NON\_TER 1

FT NON\_TER 272

SQ SEQUENCE 272 AA; 31647 MW; DFB1FFBAFDFFCB59 CRC64;

## Query Match

72.3%; Score 34; DB 7; Length 272;

Best Local Similarity 50.0%; Pred. No. 7.8;

Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10  
||| :||

Db 160 GPCXDSLLRY 169

## RESULT 13

Q860A1

ID Q860A1 PRELIMINARY; PRT; 272 AA.

AC Q860A1;

DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE MHC class I antigen (Fragment).

GN MUABTL.

OS Mus abbotti (Abbott's mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10108;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=wild; TISSUE=Tail;

RX MEDLINE=22359207; PubMed=12471122;

RA Davis B.K., Cook R.G., Rich R.R., Rodgers J.R.;

RT "Hyperconservation of the putative antigen recognition site of the MHC  
class I-b molecule TL in the subfamily Murinae: evidence that thymus  
leukemia antigen is an ancient mammalian gene.";

RL J. Immunol. 169:6890-6899(2002).

DR EMBL; AY144133; AAP23909.1; -.

DR EMBL; AY144132; AAP23909.1; JOINED.

DR GO:0016020; C:membrane; IEA.

DR GO:0005489; F:electron transporter activity; IEA.

DR GO:0006118; P:electron transport; IEA.

DR GO:0006955; P:immune response; IEA.

DR InterPro; IPR000345; CytC\_heme\_BS.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig\_c1.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR001039; MHC\_I.

DR Pfam; PF00047; Ig; 1.

DR Pfam; PF00129; MHC\_I; 1.

DR PRINTS; PR01638; MHCCLASSI.

DR ProDom; PD000050; MHC\_I; 1.

DR SMART; SM00407; IGc1; 1.

DR PROSITE; PS00190; CYTOCHROME\_C; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.

DR PROSITE; PS00290; IG\_MHC; 1.

FT NON\_TER 1

FT NON\_TER 272

SQ SEQUENCE 272 AA; 31380 MW; C8E5D1FD7028AB1A CRC64;

## Query Match

72.3%; Score 34; DB 7; Length 272;

Best Local Similarity 50.0%; Pred. No. 7.8;

Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPCXXXFIRY 10

||| :||

Db 160 GPCXDSLLRY 169

## RESULT 14

Q85ZZ9

ID Q85ZZ9 PRELIMINARY; PRT; 272 AA.

AC Q85ZZ9;

DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE MHC class I antigen (Fragment).

GN MUCATL.

OS Mus caroli (Wild mouse) (Ricefield mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10089;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Tail;

MEDLINE=22359207; PubMed=12471122;  
Davis B.K., Cook R.G., Rich R.R., Rodgers J.R.;  
"Hyperconservation of the putative antigen recognition site of the MHC  
class I-b molecule TL in the subfamily Murinae: evidence that thymus  
leukemia antigen is an ancient mammalian gene.";  
J. Immunol. 169:6890-6899(2002).  
EMBL; AY144139; AAP23912.1; -.  
EMBL; AY144138; AAP23912.1; JOINED.  
GO; GO:0016020; C:membrane; IEA.  
GO; GO:0006955; P:immune response; IEA.  
InterPro; IPR007110; Ig-like.  
InterPro; IPR003597; Ig cl.  
InterPro; IPR003006; Ig\_MHC.  
InterPro; IPR001039; MHC\_I.  
Pfam; PF00047; Ig; 1.  
PRINTS; PR01638; MHCCLASSI.  
ProDom; PD000050; MHC\_I; 1.  
SMART; SM00407; IGcl; 1.  
PROSITE; PS50835; IG\_LIKE; 1.  
PROSITE; PS00290; IG\_MHC; 1.  
NON\_TER 1  
NON\_TER 272 272  
SEQUENCE 272 AA; 31592 MW; 2FB1FFBFCFDD5914 CRC64;

Query Match 72.3%; Score 34; DB 7; Length 272;  
Best Local Similarity 50.0%; Pred. No. 7.8;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GPCXXXFYRY 10  
Db 160 GPCKDSLRY 169

JLT 15  
Z6  
Q85Z26 PRELIMINARY; PRT; 272 AA.  
Q85Z26;  
01-JUN-2003 (Tremblrel. 24, Created)  
01-JUN-2003 (Tremblrel. 24, Last sequence update)  
01-OCT-2003 (Tremblrel. 25, Last annotation update)  
MHC class I antigen (Fragment).  
MUDUTL.  
Mus dunni (Pygmy mouse) (Mus terricolor).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10110;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=wild; TISSUE=Tail;  
MEDLINE=22359207; PubMed=12471122;  
Davis B.K., Cook R.G., Rich R.R., Rodgers J.R.;  
"Hyperconservation of the putative antigen recognition site of the MHC  
class I-b molecule TL in the subfamily Murinae: evidence that thymus  
leukemia antigen is an ancient mammalian gene.";  
J. Immunol. 169:6890-6899(2002).  
EMBL; AY144145; AAP23915.1; -.  
EMBL; AY144144; AAP23915.1; JOINED.  
GO; GO:0016020; C:membrane; IEA.  
GO; GO:0005489; F:electron transporter activity; IEA.  
GO; GO:0006118; P:electron transport; IEA.  
GO; GO:0006955; P:immune response; IEA.  
InterPro; IPR000345; CytC\_heme\_BS.  
InterPro; IPR007110; Ig-like.  
InterPro; IPR003597; Ig cl.  
InterPro; IPR003006; Ig\_MHC.  
InterPro; IPR001039; MHC\_I.  
Pfam; PF00047; Ig; 1.  
PRINTS; PR01638; MHCCLASSI.  
ProDom; PD000050; MHC\_I; 1.  
SMART; SM00407; IGcl; 1.  
PROSITE; PS00190; CYTOCHROME\_C; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
FT NON\_TER 1  
FT NON\_TER 272 272  
SQ SEQUENCE 272 AA; 31489 MW; 0C8CB21EA4CD1F77 CRC64;  
Query Match 72.3%; Score 34; DB 7; Length 272;  
Best Local Similarity 50.0%; Pred. No. 7.8;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GPCXXXFYRY 10  
Db 160 GPCKDSLRY 169

Search completed: August 11, 2004, 13:09:24  
Job time : 32.9048 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 11, 2004, 13:00:42 ; Search time 50.2857 Seconds  
(without alignments)  
61.807 Million cell updates/sec

le: US-10-661-207-2

fect score: 50

ence: 1 XGCXXIYXPVC 11

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 1586107 seqs, 282547505 residues

al number of hits satisfying chosen parameters: 1586107

imum DB seq length: 0  
imum DB seq length: 2000000000

:-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

ibase : A Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

alt no.	Score	Match	Length	DB	ID	Description
1	44	88.0	50	1	AAP90597	Aap90597 Sequence
2	44	88.0	55	1	AAP90596	Aap90596 Sequence
3	44	88.0	55	1	AAP90595	Aap90595 Sequence
4	44	88.0	56	1	AAP82648	Aap82648 Pancreati
5	44	88.0	56	1	AAP82613	Aap82613 Pancreati
6	44	88.0	56	1	AAP82650	Aap82650 Pancreati
7	44	88.0	56	1	AAP82649	Aap82649 Pancreati
8	44	88.0	56	1	AAP82653	Aap82653 Pancreati
9	44	88.0	56	1	AAP82611	Aap82611 Pancreati
10	44	88.0	56	1	AAP82618	Aap82618 Pancreati
11	44	88.0	56	1	AAP82652	Aap82652 Pancreati
12	44	88.0	56	1	AAP82651	Aap82651 Pancreati
13	44	88.0	56	1	AAP82612	Aap82612 Pancreati
14	44	88.0	56	1	AAP82623	Aap82623 Pancreati
15	44	88.0	56	1	AAP80003	Aap80003 Pancreati
16	44	88.0	56	1	AAP90594	Aap90594 Sequence
17	44	88.0	56	2	AAR14631	Aar14631 Pancreas
18	44	88.0	56	2	AAR14632	Aar14632 Pancreas
19	44	88.0	56	2	AAR22622	Aar22622 PTISI enco
20	44	88.0	56	2	AAR65477	Aar65477 Pancreati
21	44	88.0	57	1	AAP81182	Aap81182 Human pan
22	44	88.0	57	1	AAP94152	Aap94152 Sequence
23	44	88.0	57	2	AAR03728	Aar03728 Modified
24	44	88.0	57	2	AAR03727	Aar03727 Modified
25	44	88.0	57	2	AAR14629	Aar14629 Pancreas

26	44	88.0	75	2	AAW26361	Aaw26361 Human pan
27	44	88.0	77	1	AAP94153	Aap94153 Sequence
28	44	88.0	77	2	AAR14630	Aar14630 Pancreas
29	44	88.0	79	1	AAP80400	Aap80400 Human pan
30	44	88.0	79	5	AAU97781	Aau97781 Human pan
31	44	88.0	89	2	AAR65482	Aar65482 Fusion pr
32	44	88.0	94	3	AAB54299	Aab54299 Human pan
33	44	88.0	94	4	ABG27312	Abg27312 Novel hum
34	44	88.0	107	2	AAR92296	Aar92296 PSKAN8 pr
35	44	88.0	107	4	ABG27311	Abg27311 Novel hum
36	43	86.0	11	6	ABU09384	Abu09384 Consensus
37	42	84.0	48	6	ABP57288	Abp57288 Kazal typ
38	42	84.0	48	6	ABP57290	Abp57290 Kazal typ
39	42	84.0	48	6	ABP57291	Abp57291 Kazal typ
40	42	84.0	61	2	AAR43918	Aar43918 Mutant mo
41	42	84.0	61	2	AAR43917	Aar43917 Mutant mo
42	42	84.0	62	2	AAR35229	Aar35229 Monitor p
43	38	76.0	45	6	ABP57286	Abp57286 Kazal typ
44	38	76.0	50	6	ABP57287	Abp57287 Kazal typ
45	38	76.0	56	1	AAP82622	Aap82622 Pancreati

ALIGNMENTS

RESULT 1  
AAP90597  
ID AAP90597 standard; protein; 50 AA.  
XX  
AC AAP90597;  
XX  
DT 25-MAR-2003 (revised)  
DT 11-JUN-1989 (first entry)  
XX  
DE Sequence of human pancreatic secretory trypsin inhibitor (PSTI).  
XX  
KW Human pancreatic secretory trypsin inhibitor; PSTI.  
XX  
OS Homo sapiens.  
XX  
PN EP300459-A.  
XX  
PD 25-JAN-1989.  
XX  
PF 20-JUL-1988; 88EP-00111704.  
XX  
PR 23-JUL-1987; 87JP-00184556.  
XX

(MOCH ) MOCHIDA PHARM CO LTD.  
(MORP ) MORISHITA PHARM CO LTD.

Nobuhara M, Kanamori T, Ogino H, Mochida E;

WPI; 1989-025643/04.

Human pancreatic secretory trypsin inhibitor - obt'd. by recombinant DNA  
techniques and free from other proteins of human origin.

Claim 20; Page 20; 30pp; English.

This sequence is produced by recombinant DNA techniques. A DNA sequence  
encoding this is inserted into a vector alongside promoter, SD and signal  
sequences and E.coli cells (pref. cell wall lipoprotein-deleted mutant  
strains) transformed. The resulting product is free of any human protein  
contaminants. It may be used in monoclonal antibody production and in  
diagnosis of human pancreatic secretory inhibitor (PSTI)-related diseases  
(Updated on 25-MAR-2003 to correct PA field.)

Sequence 50 AA;

Query Match 88.0%; Score 44; DB 1; Length 50;  
Best Local Similarity 70.0%; Pred. No. 0.57;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
||| |||  
Db 10 GCTKIYDPVC 19

RESULT 2

AAP90596  
ID AAP90596 standard; protein; 55 AA.

XX  
AC AAP90596;

XX  
DT 25-MAR-2003 (revised)

DT 11-JUN-1989 (first entry)

XX  
DE Sequence of human pancreatic secretory trypsin inhibitor (PSTI).

XX  
KW Human pancreatic secretory trypsin inhibitor; PSTI.

XX  
OS Homo sapiens.

XX  
PN EP300459-A.

XX  
PD 25-JAN-1989.

XX  
PF 20-JUL-1988; 88EP-00111704.

XX  
PR 23-JUL-1987; 87JP-00184556.

XX  
PA (MOCH ) MOCHIDA PHARM CO LTD.

PA (MORP ) MORISHITA PHARM CO LTD.

XX  
PI Nobuhara M, Kanamori T, Ogino H, Mochida E;

XX  
DR WPI; 1989-025643/04.

XX  
PT Human pancreatic secretory trypsin inhibitor - obtd. by recombinant DNA  
techniques and free from other proteins of human origin.

XX  
PS Claim 20; Page 20; 30pp; English.

XX  
CC This sequence is produced by recombinant DNA techniques. A DNA sequence  
encoding this is inserted into a vector alongside promoter, SD and signal  
sequences and E.coli cells (pref. cell wall lipoprotein-deleted mutant  
strains) transformed. The resulting product is free of any human protein  
contaminants. It may be used in monoclonal antibody production and in  
diagnosis of human pancreatic secretory inhibitor (PSTI)-related diseases  
(Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 55 AA;

Query Match 88.0%; Score 44; DB 1; Length 55;  
Best Local Similarity 70.0%; Pred. No. 0.62;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
||| |||  
Db 15 GCTKIYDPVC 24

RESULT 3

AAP90595

ID AAP90595 standard; protein; 55 AA.

XX  
AC AAP90595;

XX  
DT 25-MAR-2003 (revised)

DT 11-JUN-1989 (first entry)

XX  
DE Sequence of human pancreatic secretory trypsin inhibitor (PSTI).

XX  
KW Human pancreatic secretory trypsin inhibitor; PSTI.

OS Homo sapiens.

XX  
PN EP300459-A.

XX  
PD 25-JAN-1989.

XX  
PF 20-JUL-1988; 88EP-00111704.

XX  
PR 23-JUL-1987; 87JP-00184556.

XX  
PA (MOCH ) MOCHIDA PHARM CO LTD.

PA (MORP ) MORISHITA PHARM CO LTD.

XX  
PI Nobuhara M, Kanamori T, Ogino H, Mochida E;

XX  
DR WPI; 1989-025643/04.

XX  
PT Human pancreatic secretory trypsin inhibitor - obtd. by recombinant DNA  
techniques and free from other proteins of human origin.

XX  
PS Claim 20; Page 20; 30pp; English.

XX  
CC This sequence is produced by recombinant DNA techniques. A DNA sequence  
encoding this is inserted into a vector alongside promoter, SD and signal  
sequences and E.coli cells (pref. cell wall lipoprotein-deleted mutant  
strains) transformed. The resulting product is free of any human protein  
contaminants. It may be used in monoclonal antibody production and in  
diagnosis of human pancreatic secretory inhibitor (PSTI)-related diseases  
(Updated on 25-MAR-2003 to correct PA field.)

XX  
SQ Sequence 55 AA;

Query Match 88.0%; Score 44; DB 1; Length 55;  
Best Local Similarity 70.0%; Pred. No. 0.62;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
||| |||  
Db 15 GCTKIYDPVC 24

RESULT 4

AAP82648

ID AAP82648 standard; protein; 56 AA.

XX  
AC AAP82648;

XX  
DT 25-MAR-2003 (revised)

DT 06-NOV-1990 (first entry)

XX  
DE Pancreatic secretory trypsin inhibitor, PSTI 18 (Ile18).

XX  
KW PSTI 18; pancreatic secretory trypsin inhibitor; protease;  
leukocyte elastase.

XX  
OS Synthetic.

XX  
FH Key Location/Qualifiers

FT Misc-difference 18 /label= site-directed mutn.  
FT /note= "Lys>Ile"

XX  
PN GB2199582-A.

XX  
PD 13-JUL-1988.

XX  
PF 07-JAN-1987; 87GB-00000204.

XX  
PR 07-JAN-1987; 87GB-00000204.

XX  
PA (FARB ) BAYER AG.

XX  
PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;

WPI; 1988-192315/28.  
N-PSDB; AAN82228.  
Analogues of pancreatic secretory trypsin inhibitor - are protease inhibitors with high specificity for leukocyte elastase.  
Claim 8; Page 62; 74pp; English.  
The protein was produced from a DNA sequence constructed from 25 oligonucleotides and is one of 24 analogues of PSTI. A master gene, encoding PSTI 0, (having essentially the same sequence as reported by L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template for the construction of novel genes. See also AAP82611-23 and AAP82649-54. (Updated on 25-MAR-2003 to correct PI field.)  
Sequence 56 AA;  
Query Match 88.0%; Score 44; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.63;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
2 GCXXIYXPVC 11  
||| |||  
15 GCTIYNPVC 24  
JLT 5  
32613  
AAP82613 standard; protein; 56 AA.  
AAP82613;  
25-MAR-2003 (revised)  
06-NOV-1990 (first entry)  
Pancreatic secretory trypsin inhibitor, PSTI 3 (Tyr18,Glu19,Arg21).  
PSTI 3; pancreatic secretory trypsin inhibitor; protease;  
leukocyte elastase.  
Synthetic.  
Key Location/Qualifiers  
Misc-difference 18 /label= site-directed mutn.  
/note= "Lys>Tyr"  
Misc-difference 19 /label= site-directed mutn.  
/note= "Ile>Glu"  
Misc-difference 21 /label= site-directed mutn.  
/note= "Asn>Arg"  
GB2199582-A.  
13-JUL-1988.  
07-JAN-1987; 87GB-00000204.  
07-JAN-1987; 87GB-00000204.  
(FARB ) BAYER AG.  
Collins D, Blocker H., Frank R, Maywald F, Fritz H, Bruns W;  
WPI; 1988-192315/28.  
N-PSDB; AAN82218.  
Analogues of pancreatic secretory trypsin inhibitor - are protease inhibitors with high specificity for leukocyte elastase.  
Claim 8; Page 62; 74pp; English.

XX The protein was produced from a DNA sequence constructed from 25 oligonucleotides and is one of 24 analogues of PSTI. A master gene, encoding PSTI 0, (having essentially the same sequence as reported by L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template for the construction of novel genes. See also AAP82611-23 and AAP82648-54. (Updated on 25-MAR-2003 to correct PI field.)  
XX Sequence 56 AA;  
SQ Query Match 88.0%; Score 44; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.63;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GCXXIYXPVC 11  
||| |||  
DB 15 GCTIYRPVC 24  
RESULT 6  
AAP82650  
ID AAP82650 standard; protein; 56 AA.  
XX AAP82650;  
AC AAP82650;  
XX 25-MAR-2003 (revised)  
DT 06-NOV-1990 (first entry)  
XX Pancreatic secretory trypsin inhibitor, PSTI 20 (Ile18,Asp21,Asn29).  
DE PSTI 20; pancreatic secretory trypsin inhibitor; protease;  
XX leukocyte elastase.  
KW Synthetic.  
XX Key Location/Qualifiers  
FT Misc-difference 18 /label= site-directed mutn.  
/note= "Lys>Val"  
FT Misc-difference 21 /label= site-directed mutn.  
/note= "Asn>Asp"  
FT Misc-difference 29 /label= site-directed mutn.  
/note= "Asp>Asn"  
FT GB2199582-A.  
XX 13-JUL-1988.  
PD 07-JAN-1987; 87GB-00000204.  
XX 07-JAN-1987; 87GB-00000204.  
PR (FARB ) BAYER AG.  
XX Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;  
XX WPI; 1988-192315/28.  
DR N-PSDB; AAN82230.  
DR Analogues of pancreatic secretory trypsin inhibitor - are protease inhibitors with high specificity for leukocyte elastase.  
XX Claim 8; Page 62; 74pp; English.  
XX The protein was produced from a DNA sequence constructed from 25 oligonucleotides and is one of 24 analogues of PSTI. A master gene, encoding PSTI 0, (having essentially the same sequence as reported by L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template for the construction of novel genes. See also AAP82611-23 and AAP82648-54. (Updated on 25-MAR-2003 to correct PI field.)  
XX

```
SQ      Sequence 56 AA;
Query Match      88.0%; Score 44; DB 1; Length 56;
Best Local Similarity 70.0%; Pred. No. 0.63;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GCXXIYXPVC 11
      |||||
Db      15 GCTLIYDPVC 24

RESULT 7
AAP82649
ID AAP82649 standard; protein; 56 AA.
XX
AC AAP82649;
XX
DT 25-MAR-2003 (revised)
DT 06-NOV-1990 (first entry)
XX
DE Pancreatic secretory trypsin inhibitor, PSTI 19 (Val18,Asp21,Asn29).
XX
KW PSTI 19; pancreatic secretory trypsin inhibitor; protease;
KW leukocyte elastase.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 18 /label= site-directed mutn.
FT /note= "Lys>Val"
FT Misc-difference 21 /label= site-directed mutn.
FT /note= "Asn>Asp"
FT Misc-difference 29 /label= site-directed mutn.
FT /note= "Asp>Asn"
XX
PN GB2199582-A.
XX
PD 13-JUL-1988.
XX
PF 07-JAN-1987; 87GB-00000204.
XX
PR 07-JAN-1987; 87GB-00000204.
XX
PA (FARB ) BAYER AG.
XX
PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;
XX
DR WPI; 1988-192315/28.
DR N-PSDB; AAN82229.
XX
PT Analogues of pancreatic secretory trypsin inhibitor - are protease
PT inhibitors with high specificity for leukocyte elastase.
PS Claim 8; Page 62; 74pp; English.
XX
CC The protein was produced from a DNA sequence constructed from 25
CC oligonucleotides and is one of 24 analogues of PSTI. A master gene,
CC encoding PSTI 0, (having essentially the same sequence as reported by
CC L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template
CC for the construction of novel genes. See also AAP82611-23 and AAP82648-
CC 54. (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence 56 AA;
Query Match      88.0%; Score 44; DB 1; Length 56;
Best Local Similarity 70.0%; Pred. No. 0.63;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GCXXIYXPVC 11
      |||||
Db      15 GCTLIYDPVC 24

RESULT 8
AAP82653
ID AAP82653 standard; protein; 56 AA.
XX
AC AAP82653;
XX
DT 25-MAR-2003 (revised)
DT 06-NOV-1990 (first entry)
XX
DE Pancreatic secretory trypsin inhibitor, PSTI 23 (Phe18).
XX
KW PSTI 23; pancreatic secretory trypsin inhibitor; protease;
KW leukocyte elastase.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 18 /label= site-directed mutn.
FT /note= "Lys>Phe"
XX
PN GB2199582-A.
XX
PD 13-JUL-1988.
XX
PF 07-JAN-1987; 87GB-00000204.
XX
PR 07-JAN-1987; 87GB-00000204.
XX
PA (FARB ) BAYER AG.
XX
PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;
XX
DR WPI; 1988-192315/28.
DR N-PSDB; AAN82233.
XX
PT Analogues of pancreatic secretory trypsin inhibitor - are protease
PT inhibitors with high specificity for leukocyte elastase.
PS Claim 8; Page 62; 74pp; English.
XX
CC The protein was produced from a DNA sequence constructed from 25
CC oligonucleotides and is one of 24 analogues of PSTI. A master gene,
CC encoding PSTI 0, (having essentially the same sequence as reported by
CC L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template
CC for the construction of novel genes. See also AAP82611-23 and AAP82648-
CC 54. (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence 56 AA;
Query Match      88.0%; Score 44; DB 1; Length 56;
Best Local Similarity 70.0%; Pred. No. 0.63;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GCXXIYXPVC 11
      |||||
Db      15 GCTLIYDPVC 24

RESULT 9
AAP82611
ID AAP82611 standard; protein; 56 AA.
XX
AC AAP82611;
XX
DT 06-NOV-1990 (first entry)
XX
DE Pancreatic secretory trypsin inhibitor, PSTI 1 (Leu18).
XX
KW PSTI 1; pancreatic secretory trypsin inhibitor; protease;
```

```
Db      15 GCTVIYDPVC 24

RESULT 8
AAP82653
ID AAP82653 standard; protein; 56 AA.
XX
AC AAP82653;
XX
DT 25-MAR-2003 (revised)
DT 06-NOV-1990 (first entry)
XX
DE Pancreatic secretory trypsin inhibitor, PSTI 23 (Phe18).
XX
KW PSTI 23; pancreatic secretory trypsin inhibitor; protease;
KW leukocyte elastase.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 18 /label= site-directed mutn.
FT /note= "Lys>Phe"
XX
PN GB2199582-A.
XX
PD 13-JUL-1988.
XX
PF 07-JAN-1987; 87GB-00000204.
XX
PR 07-JAN-1987; 87GB-00000204.
XX
PA (FARB ) BAYER AG.
XX
PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;
XX
DR WPI; 1988-192315/28.
DR N-PSDB; AAN82233.
XX
PT Analogues of pancreatic secretory trypsin inhibitor - are protease
PT inhibitors with high specificity for leukocyte elastase.
PS Claim 8; Page 62; 74pp; English.
XX
CC The protein was produced from a DNA sequence constructed from 25
CC oligonucleotides and is one of 24 analogues of PSTI. A master gene,
CC encoding PSTI 0, (having essentially the same sequence as reported by
CC L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template
CC for the construction of novel genes. See also AAP82611-23 and AAP82648-
CC 54. (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence 56 AA;
Query Match      88.0%; Score 44; DB 1; Length 56;
Best Local Similarity 70.0%; Pred. No. 0.63;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GCXXIYXPVC 11
      |||||
Db      15 GCTLIYDPVC 24

RESULT 9
AAP82611
ID AAP82611 standard; protein; 56 AA.
XX
AC AAP82611;
XX
DT 06-NOV-1990 (first entry)
XX
DE Pancreatic secretory trypsin inhibitor, PSTI 1 (Leu18).
XX
KW PSTI 1; pancreatic secretory trypsin inhibitor; protease;
```



leukocyte elastase.  
 Synthetic.  
 Key Location/Qualifiers  
 Misc-difference 18 /label= site-directed mutn.  
 /note= "Lys>Leu"  
 GB2199582-A.  
 13-JUL-1988.  
 07-JAN-1987; 87GB-00000204.  
 07-JAN-1987; 87GB-00000204.  
 (FARB ) BAYER AG.  
 Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;  
 WPI; 1988-192315/28.  
 N-PSDB; AAN82216.  
 Analogues of pancreatic secretory trypsin inhibitor - are protease inhibitors with high specificity for leukocyte elastase.  
 Claim 8; Page 62; 74pp; English.  
 The protein was produced from a DNA sequence constructed from 25 oligonucleotides and is one of 24 analogues of PSTI. A master gene, encoding PSTI 0, (having essentially the same sequence as reported by L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template for the construction of novel genes. See also AAP82612-23 and AAP82648-54  
 Sequence 56 AA;  
 Query Match 88.0%; Score 44; DB 1; Length 56;  
 Best Local Similarity 70.0%; Pred. No. 0.63;  
 Matches 7; Conservative 0; Mismatches 3; Indels 3; Gaps 0;  
 QY 2 GCXXIYXPVC 11  
 DB 15 GCTLIYNPVC 24  
 JLT 10  
 32618  
 AAP82618 standard; protein; 56 AA.  
 AAP82618;  
 25-MAR-2003 (revised)  
 06-NOV-1990 (first entry)  
 Pancreatic secretory trypsin inhibitor, PSTI 7 (Leu18,Arg21).  
 PSTI 7; pancreatic secretory trypsin inhibitor; protease;  
 leukocyte elastase.  
 Synthetic.  
 Key Location/Qualifiers  
 Misc-difference 18 /label= site-directed mutn.  
 /note= "Lys>Leu"  
 Misc-difference 21 /label= site-directed mutn.  
 /note= "Asn>Arg"  
 GB2199582-A.  
 13-JUL-1988.

XX 07-JAN-1987; 87GB-00000204.  
 PF XX  
 PR 07-JAN-1987; 87GB-00000204.  
 XX (FARB ) BAYER AG.  
 PA XX  
 PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;  
 XX WPI; 1988-192315/28.  
 DR N-PSDB; AAN82222.  
 XX Analogues of pancreatic secretory trypsin inhibitor - are protease inhibitors with high specificity for leukocyte elastase.  
 PT Claim 8; Page 62; 74pp; English.  
 XX The protein was produced from a DNA sequence constructed from 25 oligonucleotides and is one of 24 analogues of PSTI. A master gene, encoding PSTI 0, (having essentially the same sequence as reported by L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template for the construction of novel genes. See also AAP82611-23 and AAP82648-54. (Updated on 25-MAR-2003 to correct PI field.)  
 XX Sequence 56 AA;  
 SQ Query Match 88.0%; Score 44; DB 1; Length 56;  
 Best Local Similarity 70.0%; Pred. No. 0.63;  
 Matches 7; Conservative 0; Mismatches 3; Indels 3; Gaps 0;  
 QY 2 GCXXIYXPVC 11  
 DB 15 GCTLIYRPVC 24  
 RESULT 11  
 AAP82652  
 ID AAP82652 standard; protein; 56 AA.  
 XX AC AAP82652;  
 XX 25-MAR-2003 (revised)  
 DT 06-NOV-1990 (first entry)  
 XX Pancreatic secretory trypsin inhibitor, PSTI 22 (Tyr18).  
 DE PSTI 22; pancreatic secretory trypsin inhibitor; protease;  
 KW leukocyte elastase.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 FH Misc-difference 18 /label= site-directed mutn.  
 FT /note= "Lys>Tyr"  
 FT GB2199582-A.  
 XX PN  
 XX 13-JUL-1988.  
 PD 07-JAN-1987; 87GB-00000204.  
 PF 07-JAN-1987; 87GB-00000204.  
 XX (FARB ) BAYER AG.  
 PR Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;  
 PA WPI; 1988-192315/28.  
 XX N-PSDB; AAN82232.  
 DR Analogues of pancreatic secretory trypsin inhibitor - are protease inhibitors with high specificity for leukocyte elastase.  
 PT

XX Claim 8; Page 62; 74pp; English.

XX The protein was produced from a DNA sequence constructed from 25

CC oligonucleotides and is one of 24 analogues of PSTI. A master gene,

CC encoding PSTI 0, (having essentially the same sequence as reported by

CC L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template

CC for the construction of novel genes. See also AAP82611-23 and AAP82648-

CC 54. (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 56 AA;

Query Match 88.0%; Score 44; DB 1; Length 56;

Best Local Similarity 70.0%; Pred. No. 0.63;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11

Db 15 GCTIYINPVC 24

RESULT 12

AAP82651

ID AAP82651 standard; protein; 56 AA.

XX

AC AAP82651;

XX

DT 25-MAR-2003 (revised)

DT 06-NOV-1990 (first entry)

XX

DE Pancreatic secretory trypsin inhibitor, PSTI 21 (Ile18,Glu19,Arg21).

XX

KW PSTI 21; pancreatic secretory trypsin inhibitor; protease;

KW leukocyte elastase.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 18 /label= site-directed mutn.

FT /note= "Lys>Val"

FT Misc-difference 19 /label= site-directed mutn.

FT /note= "Ile>Glu"

FT Misc-difference 21 /label= site-directed mutn.

FT /note= "Asn>Arg"

XX

PN GB2199582-A.

XX

PD 13-JUL-1988.

XX

PF 07-JAN-1987; 87GB-00000204.

XX

PR 07-JAN-1987; 87GB-00000204.

XX

PA (FARB ) BAYER AG.

XX

PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;

XX

WPI; 1988-192315/28.

DR N-PSDB; AAN82231.

XX

XX Analogues of pancreatic secretory trypsin inhibitor - are protease

PT inhibitors with high specificity for leukocyte elastase.

XX

PS Claim 8; Page 62; 74pp; English.

XX

CC The protein was produced from a DNA sequence constructed from 25

CC oligonucleotides and is one of 24 analogues of PSTI. A master gene,

CC encoding PSTI 0, (having essentially the same sequence as reported by

CC L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template

CC for the construction of novel genes. See also AAP82611-23 and AAP82648-

CC 54. (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 56 AA;

Query Match 88.0%; Score 44; DB 1; Length 56;

Best Local Similarity 70.0%; Pred. No. 0.63;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11

Db 15 GCTLIYDPVC 24

RESULT 13

AAP82612

ID AAP82612 standard; protein; 56 AA.

XX

AC AAP82612;

XX

DT 25-MAR-2003 (revised)

DT 06-NOV-1990 (first entry)

XX

DE Pancreatic secretory trypsin inhibitor, PSTI 2 (Leu18,Asp21,Asn29).

XX

KW PSTI 2; pancreatic secretory trypsin inhibitor; protease;

KW leukocyte elastase.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 18 /label= site-directed mutn.

FT /note= "Lys>Leu"

FT Misc-difference 21 /label= site-directed mutn.

FT /note= "Asn>Asp"

FT Misc-difference 29 /label= site-directed mutn.

FT /note= "Asp>Asn"

XX

PN GB2199582-A.

XX

PD 13-JUL-1988.

XX

PF 07-JAN-1987; 87GB-00000204.

XX

PR 07-JAN-1987; 87GB-00000204.

XX

PA (FARB ) BAYER AG.

XX

PI Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;

XX

WPI; 1988-192315/28.

DR N-PSDB; AAN82217.

XX

XX Analogues of pancreatic secretory trypsin inhibitor - are protease

PT inhibitors with high specificity for leukocyte elastase.

XX

PS Claim 8; Page 62; 74pp; English.

XX

CC The protein was produced from a DNA sequence constructed from 25

CC oligonucleotides and is one of 24 analogues of PSTI. A master gene,

CC encoding PSTI 0, (having essentially the same sequence as reported by

CC L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template

CC for the construction of novel genes. See also AAP82611-23 and AAP82648-

CC 54. (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 56 AA;

Query Match 88.0%; Score 44; DB 1; Length 56;

Best Local Similarity 70.0%; Pred. No. 0.63;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 GCXXIYXPVC 11  
|||  
15 GCTLIYDPVC 24

HLT 14

12623  
AAP82623 standard; protein; 56 AA.

AAP82623;

25-MAR-2003 (revised)  
06-NOV-1990 (first entry)

Pancreatic secretory trypsin inhibitor, PSTI 17 (Val18).

PSTI 17; pancreatic secretory trypsin inhibitor; protease;  
leukocyte elastase.

Synthetic.

Key Location/Qualifiers  
Misc-difference 18  
/label= site-directed mutn.  
/note= "Lys>Val"

GB2199582-A.

13-JUL-1988.

07-JAN-1987; 87GB-00000204.

07-JAN-1987; 87GB-00000204.

(FARB ) BAYER AG.

Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;

WPI; 1988-192315/28.

N-PSDB; AAN82227.

Analogues of pancreatic secretory trypsin inhibitor - are protease  
inhibitors with high specificity for leukocyte elastase.

Claim 8; Page 62; 74pp; English.

The protein was produced from a DNA sequence constructed from 25  
oligonucleotides and is one of 24 analogues of PSTI. A master gene,  
encoding PSTI 0, (having essentially the same sequence as reported by  
L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as a template  
for the construction of novel genes. See also AAP82611-22 and AAP82648-  
54. (Updated on 25-MAR-2003 to correct PI field.)

Sequence 56 AA;

Query Match 88.0%; Score 44; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.63;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 GCXXIYXPVC 11  
|||  
15 GCTVIYNPVC 24

HLT 15  
10003

AAP80003 standard; protein; 56 AA.

AAP80003;

06-NOV-1990 (first entry)

Pancreatic secretory trypsin inhibitor, PSTI 0.

XX PSTI 0; pancreatic secretory trypsin inhibitor; protease;  
KW leukocyte elastase.  
XX  
OS Synthetic.  
XX GB2199582-A.  
PN  
XX 13-JUL-1988.  
PD  
XX 07-JAN-1987; 87GB-00000204.  
PF  
XX 07-JAN-1987; 87GB-00000204.  
PR  
XX (FARB ) BAYER AG.  
PA  
XX Collins D, Blocker H, Frank R, Maywald F, Fritz H, Bruns W;  
PI  
XX WPI; 1988-192315/28.  
DR  
XX N-PSDB; AAN80030.  
DR  
XX Analogues of pancreatic secretory trypsin inhibitor - are protease  
PT inhibitors with high specificity for leukocyte elastase.  
PT  
XX Disclosure; Page 7; 74pp; English.  
PS  
XX The protein was produced from a DNA sequence constructed from 25  
CC oligonucleotides and has the same sequence as the natural PSTI 0. A  
CC master gene, encoding PSTI 0, (having essentially the same sequence as  
CC reported by L.J. Greene, 1976, Methods Enzymol. 45, 813-825) was used as  
CC a template for the construction of novel genes. See also AAP82611-23 and  
CC AAP82648-54  
XX  
SQ Sequence 56 AA;

Query Match 88.0%; Score 44; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.63;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GCXXIYXPVC 11  
|||  
Db 15 GCTKIYNPVC 24

Search completed: August 11, 2004, 13:07:40  
Job time : 51.2857 secs

**BEST AVAILABLE COPY**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 11, 2004, 13:04:58 ; Search time 12.0475 Seconds  
(without alignments)  
87.827 Million cell updates/sec

le: US-10-661-207-2

fect score: 50

ence: 1 XGCXXIYXPVC 11

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 283366 seqs, 96191526 residues

al number of hits satisfying chosen parameters: 283366

imum DB seq length: 0  
imum DB seq length: 2000000000

--processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

ibase : PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

alt No.	Score	Query		Length	DB	ID	Description
		Match					
1	44	88.0	79	1	TIHUA	pancreatic secreto	
2	42	84.0	56	1	TIBOA	pancreatic secreto	
3	42	84.0	56	1	TISHA	pancreatic secreto	
4	42	84.0	56	1	TIPG.	pancreatic secreto	
5	42	84.0	57	1	TIDGA	pancreatic secreto	
6	42	84.0	79	1	TIRT1	pancreatic secreto	
7	42	84.0	80	1	S01498	pancreatic secreto	
8	38	76.0	86	1	A34427	peptide PEC-60 pre	
9	38	76.0	472	1	A26730	ovoinhibitor precu	
10	38	76.0	1959	1	AGRT	agrin - rat	
11	37	74.0	53	2	G31438	ovomucoid, third d	
12	36	72.0	1699	2	T31340	voltage-gated sodi	
13	35	70.0	79	1	TIRT2	pancreatic secreto	
14	35	70.0	93	2	JA0128	hypothetical l1k p	
15	35	70.0	1154	2	T18525	diacylglycerol kin	
16	34	68.0	94	1	B48549	nucleic acid-bindi	
17	34	68.0	341	2	A48422	ecdysone-inducible	
18	34	68.0	746	2	G84605	hypothetical prote	
19	33	66.0	54	2	H31444	ovomucoid, third d	
20	33	66.0	772	2	H84605	hypothetical prote	
21	33	66.0	802	2	B84560	hypothetical prote	
22	33	66.0	2422	2	T18201	fatty-acid synthas	
23	32	64.0	128	2	A29349	hypothetical prote	
24	32	64.0	201	2	C85253	receptor like prot	
25	32	64.0	201	2	T49121	receptor like prot	
26	32	64.0	223	2	F82870	thymidine kinase U	
27	32	64.0	572	2	T16865	hypothetical prote	
28	32	64.0	609	2	T25120	hypothetical prote	
29	32	64.0	727	2	S42834	F40F12.5 protein -	

30	32	64.0	806	2	A53256	nuclear protein bi
31	32	64.0	1021	2	F88568	protein F40F12.5 [
32	32	64.0	1062	2	T14151	inv protein - mous
33	32	64.0	1062	2	T30255	inversin - mouse
34	32	64.0	1872	2	S64851	probable membrane
35	32	64.0	1955	1	AGCH	agrin precursor -
36	31.5	63.0	759	2	S53922	PMT6 protein - yea
37	31	62.0	48	1	S00130	elastase inhibitor
38	31	62.0	53	2	A31437	ovomucoid, third d
39	31	62.0	53	2	I31438	ovomucoid, third d
40	31	62.0	56	2	B31440	ovomucoid, third d
41	31	62.0	56	2	F31443	acrosin/trypsin in
42	31	62.0	81	1	S29820	acrosin/trypsin in
43	31	62.0	84	1	JU0152	hypothetical prote
44	31	62.0	191	2	H81803	thymidine kinase
45	31	62.0	212	2	T02888	

ALIGNMENTS

RESULT 1

TIHUA

pancreatic secretory trypsin inhibitor precursor [validated] - human

N;Alternate names: endothelial cell growth factor 2a

C;Species: Homo sapiens (man)

C;Date: 30-Nov-1980 #sequence revision 17-Feb-1994 #text change 08-Dec-2000

C;Accession: A27484; S02605; A90062; A92355; A25604; I52210; A01229

R;Horii, A.; Kobayashi, T.; Tomita, N.; Yamamoto, T.; Fukushima, S.; Murotsu, T.; Ogawa,

Biochem. Biophys. Res. Commun. 149, 635-641, 1987

A;Title: Primary structure of human pancreatic secretory trypsin inhibitor (PSTI) gene.

A;Reference number: A27484; MUID:88106485; PMID:3501289

A;Accession: A27484

A;Molecule type: DNA

A;Residues: 1-79 <HOR>

A;Cross-references: GB:M20530; GB:M18374; NID:gl90692; PIDN:AAA36522.1; PID:gl90694

R;Tomita, N.; Horii, A.; Yamamoto, T.; Ogawa, M.; Mori, T.; Matsubara, K.

FEBS Lett. 225, 113-119, 1987

A;Title: Expression of pancreatic secretory trypsin inhibitor gene in neoplastic tissues

A;Reference number: S02605; MUID:88083571; PMID:2961612

A;Accession: S02605

A;Molecule type: mRNA

A;Residues: 1-63, 'G', 65-79 <TOM>

A;Cross-references: EMBL:Y00705; NID:g35765; PIDN:CAA68697.1; PID:g35766

A;Note: the authors translated the codon TAC for residue 33 as Thr and GGT for residue 6'

R;Bartelt, D.C.; Shapanka, R.; Greene, L.J.

Arch. Biochem. Biophys. 179, 189-199, 1977

A;Title: The primary structure of the human pancreatic secretory trypsin inhibitor. Amino

A;Reference number: A90062; MUID:77133145; PMID:843082

A;Accession: A90062

A;Molecule type: protein

A;Residues: 24-43, 'N', 45-51, 'D', 53-79 <BAR>

A;Note: the inhibitor is present in multiple chromatographic forms differing in asparagin

R;Huhtala, M.L.; Pesonen, K.; Kalkkinen, N.; Stenman, U.H.

J. Biol. Chem. 257, 13713-13716, 1982

A;Title: Purification and characterization of a tumor-associated trypsin inhibitor from t

A;Reference number: A92355; MUID:83056875; PMID:7142173

A;Accession: A92355

A;Molecule type: protein

A;Residues: 24-31, 'X', 33-38, 'X', 40-43, 'N', 45-46 <HUH>

A;Note: this peptide was isolated from the urine of a patient with ovarian cancer

R;McKeehan, W.L.; Sakagami, Y.; Hoshi, H.; McKeehan, K.A.

J. Biol. Chem. 261, 5378-5383, 1986

A;Title: Two apparent human endothelial cell growth factors from human hepatoma cells are

A;Reference number: A92583; MUID:86168278; PMID:3007499

A;Accession: A25604

A;Molecule type: protein

A;Residues: 24-31, 'X', 33-38, 'X', 40-46, 'X', 48 <MCK>

R;Yamamoto, T.; Nakamura, Y.; Nishide, T.; Emi, M.; Ogawa, M.; Mori, T.; Matsubara, K.

Biochem. Biophys. Res. Commun. 132, 605-612, 1985

A;Title: Molecular cloning and nucleotide sequence of human pancreatic secretory trypsin

A;Reference number: I52210; MUID:86050645; PMID:3877508

A;Accession: I52210

A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-79 <RES>  
A:Cross-references: GB:M11949; NID:gl90687; PIDN:AAA36521.1; PID:gl90688  
C:Genetics:

Query Match	88.0%;	Score 44;	DB 1;	Length 79;
Best Local Similarity	70.0%;	Pred. No. 0.11;		
Matches 7;	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps	0;

Qy 2 GCXXIYXPVC 11  
pb 38 GCTKIYDPVC 47

## RESULT 2

TIBOA  
pancreatic secretory trypsin inhibitor - bovine  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 16-Jul-1999  
C/Accession: A01230  
R/Greene, L.J.; Bartelt, D.C.  
J. Biol. Chem. 244, 2646-2657, 1969  
A/Title: The structure of the bovine pancreatic secretory trypsin inhibitor - Kazal's in  
A/Reference number: A92046;MUID:69187206; PMID:5769997

Accession: A01250  
A:Molecule type: protein  
A:Residues: 1-56 <GRE>  
R:Guy, O.; Shapanka, R.; Greene, L.J.  
J. Biol. Chem. 246, 7740-7747, 1971  
A:Title: The structure of the bovine pancreatic secretory trypsin inhibitor-Kazal's inhibitor  
A:Reference number: A92057; MUID:72086018; PMID:5135319  
A:Contents: annotation; disulfide bonds  
R:Riqbi, M.; Greene, L.J.

J. Biol. Chem. 243, 5457-5464, 1968  
A>Title: Limited proteolysis of the bovine pancreatic secretory trypsin inhibitor at acid  
A:Reference number: A92037; MUID:69080142; PMID:5750336  
A:Contents: annotation; inhibitory site

Query Match	84.0%;	Score 42;	DB 1;	Length 56;
Best Local Similarity	70.0%;	Pred. No. 0.21;		
Matches	7;	Conservative	0;	Mismatches 3;
				Indels 0;
				Gaps 0;

Qy	2	GCXXIYXPVC	11
D/b	15	GCPRIYNPVC	24

### RESULT 3

pancreatic secretory trypsin inhibitor - sheep (tentative sequence)  
 C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
 C/Accession: A94487; A01230  
 R:/Tschesche, H.; Obermeier, R.; Hochstrasser, K.  
 unpublished results, cited by Tschesche, H., Wächter, E., Kupfer, S., Obermeier, R., Reib  
 Inhibitors, Fritz, H., and Tschesche, H., eds., pp.207-222, Walter de Gruyter, New York

A;Reference number: A94487  
A;Accession: A94487  
A;Molecule type: protein  
A;Residues: 1-56 <TSC>  
C;Superfamily: pancreatic secretory trypsin inhibitor; Kazal proteinase inhibitor hom  
C;Keywords: pancreas; serine proteinase inhibitor  
F;7-56/Domain: Kazal proteinase inhibitor homology <KPI>  
F;9-38,16-35,24-56/Disulfide bonds: #status predicted  
F;18/Inhibitory site: Arg (trypsin) #status predicted

Query Match 84.0%; Score 42; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. NO. 0.21;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	2	GCXXIYXPVC	11
Db	15	GCPRIYNPVC	24

## RESULT 4

**GOIT**

pancreatic secretory trypsin inhibitor - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 16-Jul-1999  
C;Accession: A91174; A92079; A91647; A90646; A01231  
R;Tschesche, H.; Wachter, E.  
Eur. J. Biochem. 16, 187-198, 1970  
A;Title: The structure of the porcine pancreatic secretory trypsin inhibitor I. A seq  
A;Reference number: A91174; MUID:70283430; PMID:5466061

A;Concns: FSII 1  
A;Accession: A91174  
A:Molecule type: protein

A;Residues: 1-56 <TS1>  
R;Bartelt, D.C.; Greene, L.J.  
J. Biol. Chem. 246, 2218-2229, 1971  
A;Title: The primary structure of the porcine pancreatic secretory trypsin inhibitor :  
A;Reference number: A93079;MUID:71178430; PMID:5103069

A;Accession: A92079  
A;Molecule type: protein  
A;Residues: 1-56 <BAR>  
R;Tschesche, H.; Schneider, M.; Reidel, G.; Klein, H. Hoppe-Seyler's Z. Physiol. Chem. 353, 763-764, 1972  
A;Title: Die Disulfidbruecken des sekretorischen Schweinepankreas-Trypsininhibitors u  
A;Reference number: A91654; MUID: 73001409; PMID: 4672150

A; Contents: annotation; disulfide bonds  
R; Tschesche, H.; Wachter, E.

Hoppe-Seyler's Z. Physiol. Chem. 351, 1449-1459, 1970  
A;Title: Die Primaerstruktur des spezifischen Trypsininhibitors II (Kazal-Typ) aus Schweinepankreas  
A;Reference number: A91647; MUID:71092915; PMID:5531651  
A;Contents: PSTI II  
A;Accession: A91647  
A;Molecule type: protein  
A;Residues: 5-56 <TS2>  
R;Menegatti, E.; Bortolotti, F.; Minchiotti, L.; de Marco, A.  
Biochim. Biophys. Acta 707, 50-58, 1982  
A;Title: Isolation and characterization of a new form of the porcine pancreatic secretory trypsin inhibitor (PSTI II)  
A;Reference number: A90646; MUID:83049107; PMID:7138878

Query Match 84.0%; Score 42; DB 1; Length 56;  
Best Local Similarity 70.0%;  
Matches 7; Conservative 0; Mismatches 3; Indels  
Pred. No. 0.21;

QY 2 GCXXIXPVC 11

```

15 GCPKIYNPVC 24

>Title: Molecular cloning and characterization of genes encoding rat pancreatic cholecystokinin receptor type 1A
>Reference number: S28946; MUID:93003324; PMID:1390891
>Accession: S28946
>Molecule type: DNA
>Residues: 1-79 <TS2>
>Cross-references: DDBJ:D11321; NID:g220694; PIDN:BAA01944.1; PID:g220696
R;Uda, K.I.; Ogawa, M.; Shibata, T.; Murata, A.; Mori, T.; Kikuchi, N.; Yoshida, N.; Tsui
Biol. Chem. Hoppe-Seyler 369(Suppl.), 55-61, 1988
>Title: Purification, characterization and amino-acid sequencing of two pancreatic secretory
>Reference number: S00633; MUID:89076534; PMID:3202973
>Accession: S00633
>Molecule type: protein
>Residues: 19-79 <UDA>
R;Iwai, K.; Fukuoaka, S.I.; Fushiki, T.; Tsujikawa, M.; Hirose, M.; Tsunasawa, S.; Sakiyama
J. Biol. Chem. 262, 8956-8959, 1987
>Title: Purification and sequencing of a trypsin-sensitive cholecystokinin-releasing peptide
>Reference number: A27111; MUID:87250528; PMID:3597401
>Accession: A27111
>Molecule type: protein
>Residues: 19-56,'Z',58-77,'G',79 <IWA>
R;Fukuoka, S.
Pancreas 4, 1-7, 1990
>Title: Rapid and selective cloning of monitor peptide, a novel CCK-releasing peptide, from
>Reference number: I58414
>Accession: I58414
>Status: preliminary; translated from GB/EMBL/DDBJ
>Molecule type: mRNA
>Residues: 1-79 <RES>
>Cross-references: GB:M35299; NID:g950097; PIDN:AAA74479.1; PID:g206467
>Accession: I78898
>Status: preliminary; translated from GB/EMBL/DDBJ
>Molecule type: mRNA
>Residues: 40-76 <RE2>
>Cross-references: GB:M35300; NID:g206470; PIDN:AAA41977.1; PID:g206471
C;Genetics:
>Gene: PSTI
>Introns: 19/1; 29/3; 65/2
C;Superfamily: pancreatic secretory trypsin inhibitor; Kazal proteinase inhibitor homologous
C;Keywords: pancreas; serine proteinase inhibitor
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-79/Product: pancreatic secretory trypsin inhibitor I #status experimental <MAT>
F;30-79/Domain: Kazal proteinase inhibitor homology <KPI>
F;32-61,39-58,47-79/Disulfide bonds: #status predicted
F;41/Inhibitory site: Arg (trypsin) #status predicted

Query Match      84.0%; Score 42; DB 1; Length 79;
Best Local Similarity 70.0%; Pred. No. 0.27;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GCXXIXXPVC 11
       ||| |||| 
Db      38 GCPRIYDPVC 47

RESULT 7
S01498
pancreatic secretory trypsin inhibitor precursor - mouse
N;Alternate names: prostatic secretory glycoprotein p12
C;Species: Mus musculus (house mouse)
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
>Accession: S01498; S18384
R;Mills, J.S.; Needham, M.; Parker, M.G.
EMBO J. 6, 3711-3717, 1987
>Title: A secretory protease inhibitor requires androgens for its expression in male sex
>Reference number: S01498; MUID:88111560; PMID:3428272
>Accession: S01498
>Molecule type: mRNA
>Residues: 1-80 <NIL>
>Cross-references: EMBL:X06342; NID:g53686; PIDN:CAA29648.1; PID:g53687
A;Note: the authors translated the codon ATA for residue 69 as Met
R;Lai, M.L.; Chen, S.W.; Chen, Y.H.
Arch. Biochem. Biophys. 290, 265-271, 1991

```



A;Title: Purification and characterization of a trypsin inhibitor from mouse seminal vesicle  
 A;Reference number: S18384; MUID:92027737; PMID:1929395  
 A;Accession: S18384

A;Molecule type: protein  
 A;Residues: 43-47;49-58;60-61;63-72 <LA1>  
 C;Comment: The expression of this inhibitor is dependent upon testosterone in prostate gland  
 C;Superfamily: pancreatic secretory trypsin inhibitor; Kazal proteinase inhibitor homolog  
 C;Keywords: pancreas; prostate; seminal vesicle; serine proteinase inhibitor; testis  
 F;1-23/Domain: signal sequence #status predicted <SIG>  
 F;24-80/Product: pancreatic secretory trypsin inhibitor #status predicted <MAT>  
 F;31-80/Domain: Kazal proteinase inhibitor homolog <KPI>  
 F;33-62,40-59,48-80/Disulfide bonds: #status predicted  
 F;42/Inhibitory site: Arg (trypsin) #status predicted

Query Match 84.0%; Score 42; DB 1; Length 80;  
 Best Local Similarity 70.0%; Pred. No. 0.28;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GCXIYXPVC 11  
 || || || ||  
 Db 39 GCPRIYDPVC 48

## RESULT 8

A34427 peptide PEC-60 precursor - pig  
 C;Species: Sus scrofa domestica (domestic pig)  
 C;Date: 17-Feb-1994 #sequence revision 17-Feb-1994 #text\_change 16-Jul-1999  
 C;Accession: A44041; A34427; S24956

R;Metals, M.; Cintra, A.; Solfrini, V.; Ernfor, P.; Bortolotti, F.; Morrasutti, D.G.; C  
 J; Biol. Chem. 267, 19829-19832, 1992

A;Title: Molecular cloning of PEC-60 and expression of its mRNA and peptide in the gastr  
 A;Reference number: A44041; MUID:93015834; PMID:1400298

A;Accession: A44041  
 A;Molecule type: mRNA  
 A;Residues: 1-86 <MET>

A;Cross-references: EMBL:X67109; NID:G2033; PIDN:CAA47482.1; PID:G2034  
 A;Note: sequence extracted from NCBI backbone (NCBIP:115615)  
 R;Agarberth, B.; Soederling-Barros, J.; Joernvall, H.; Chen, Z.; Oestenson, C.G.; Efendi  
 Proc. Natl. Acad. Sci. U.S.A. 86, 8590-8594, 1989

A;Title: Isolation and characterization of a 60-residue intestinal peptide structurally  
 A;Reference number: A34427; MUID:90046843; PMID:2573065  
 A;Accession: A34427

A;Molecule type: protein  
 A;Residues: 27-86 <AGE>  
 C;Comment: This peptide does not inhibit trypsin. Its biological function is unknown, al  
 C;Comment: This peptide is synthesized in duodenal goblet cells and in monocytes in bone  
 C;Superfamily: pancreatic secretory trypsin inhibitor; Kazal proteinase inhibitor homolog  
 C;Keywords: intestine; leukocyte; serine proteinase inhibitor  
 F;1-26/Domain: signal sequence #status predicted <SIG>  
 F;27-86/Product: peptide PEC-60 #status experimental <MAT>  
 F;35-86/Domain: Kazal proteinase inhibitor homolog <KPI>  
 F;37-68,46-65/Disulfide bonds: #status predicted  
 F;48/Inhibitory site: Arg (unidentified proteinase) #status predicted  
 F;54-86/Disulfide bonds: #status experimental

Query Match 76.0%; Score 38; DB 1; Length 86;  
 Best Local Similarity 66.7%; Pred. No. 1.7;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
 | || || ||  
 Db 46 CSRIYDPVC 54

## RESULT 9

A26730 ovininhibitor precursor [validated] - chicken  
 N;Contains: serine proteinase inhibitor (Kazal type)  
 C;Species: Gallus gallus (chicken)  
 C;Date: 18-Aug-2000 #sequence revision 18-Aug-2000 #text\_change 01-Sep-2000  
 C;Accession: A26730

R;Scott, M.J.; Huckaby, C.S.; Kato, I.; Kohr, W.J.; Laskowski Jr., M.; Tsai, M.J.; O'Mall

J. Biol. Chem. 262, 5899-5907, 1987

A;Title: Ovininhibitor introns specify functional domains as in the related and linked  
 A;Reference number: A26730; MUID:87194792; PMID:3571241

A;Accession: A26730

A;Molecule type: mRNA; protein

A;Residues: 1-472 <SCO>

A;Cross-references: GB:M16141; NID:G212483; PIDN:AAA48994.1; PID:G212485  
 A;Note: parts of this sequence, including the amino end of the mature protein, were d  
 C;Comment: This is a major inhibitor in blood plasma and also occurs in egg white.

C;Superfamily: chicken ovininhibitor; Kazal proteinase inhibitor homolog  
 C;Keywords: duplication; egg white; plasma; serine proteinase inhibitor

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-472/Product: ovininhibitor #status experimental <MAT>

F;26-85/Domain: Kazal proteinase inhibitor homolog <KPI1>

F;92-150/Domain: Kazal proteinase inhibitor homolog <KPI2>

F;157-216/Domain: Kazal proteinase inhibitor homolog <KPI3>

F;223-282/Domain: Kazal proteinase inhibitor homolog <KPI4>

F;289-347/Domain: Kazal proteinase inhibitor homolog <KPI5>

F;354-413/Domain: Kazal proteinase inhibitor homolog <KPI6>

F;420-472/Domain: Kazal proteinase inhibitor homolog <KPI7>

F;28-67,45-64,53-85,94-132,110-129,118-150,159-198,176-195,184-216,225-264,242-261,251

F;47/Inhibitory site: Arg (serine proteinase) #status predicted

F;112/Inhibitory site: Arg (serine proteinase) #status predicted

F;178/Inhibitory site: Arg (serine proteinase) #status predicted

F;244/Inhibitory site: Arg (serine proteinase) #status predicted

F;309/Inhibitory site: Phe (serine proteinase) #status predicted

F;375/Inhibitory site: Met (serine proteinase) #status predicted

F;434/Inhibitory site: Met (serine proteinase) #status predicted

Query Match 76.0%; Score 38; DB 1; Length 472;  
 Best Local Similarity 66.7%; Pred. No. 7.1;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
 | || || ||

Db 373 CTMIYDPVC 381

## RESULT 10

AGRT

agrin - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 17-Nov-2000

C;Accession: JH0399; A38856

R;Rupp, F.; Payan, D.G.; Magill-Solc, C.; Cowan, D.M.; Scheller, R.H.

Neuron 6, 811-823, 1991

A;Title: Structure and expression of a rat agrin.

A;Reference number: JH0399; MUID:91222570; PMID:1851019

A;Accession: JH0399

A;Molecule type: mRNA

A;Residues: 1-1779;1799-1959 <RUP>

A;Cross-references: GB:M64780; NID:G202798; PIDN:AAA40703.1; PID:G202800

A;Experimental source: embryonic spinal cord

A;Note: it is uncertain whether Met-1, Met-18, or Met-24 is the initiator

R;Rupp, F.; Oezcelik, T.; Linial, M.; Peterson, K.; Francke, U.; Scheller, R.

J. Neurosci. 12, 3535-3544, 1992

A;Title: Structure and chromosomal localization of the mammalian agrin gene.

A;Reference number: A38856; MUID:92407628; PMID:1326608

A;Accession: A38856

A;Molecule type: mRNA

A;Residues: 1780-1798 <RU2>

A;Cross-references: GB:S44194

C;Comment: This protein mediates the motor neuron-induced aggregation of acetylcholine  
 C;Comment: 90% of rat embryonic transcripts encode the variant labeled below as form 1  
 ycholine receptor clustering activity.

C;Superfamily: agrin; EGF homolog; Kazal proteinase inhibitor homolog; laminin G ref  
 C;Keywords: alternative splicing; duplication; glycoprotein; neuromuscular junction

F;1-1959/Product: agrin, form 1 #status predicted <AG1>

F;1-1787,1799-1959/Product: agrin, form 4 #status predicted <AG4>

F;1-1779,1799-1959/Product: agrin, form 3 #status predicted <AG3>

F;1-1779,1788-1959/Product: agrin, form 5 #status predicted <AG5>

F;1-1143,1153-1959/Product: agrin, form 2 #status predicted <AG2>

F;22-50/Region: hydrophobic



```
-137/Domain: Kazal proteinase inhibitor homology <KPI1>
3-212/Domain: Kazal proteinase inhibitor homology <KPI2>
6-284/Domain: Kazal proteinase inhibitor homology <KPI3>
7-356/Domain: Kazal proteinase inhibitor homology <KPI4>
1-429/Domain: Kazal proteinase inhibitor homology <KPI5>
6-494/Domain: Kazal proteinase inhibitor homology <KPI6>
1-559/Domain: Kazal proteinase inhibitor homology <KPI7>
0-542/Region: motor neuron attachment (L-R-E) motif
16-645/Domain: Kazal proteinase inhibitor homology <KPI8>
18-739/Domain: laminin-type EGF-like homology <LE1>
2-786/Domain: laminin-type EGF-like homology <LE2>
4-864/Domain: Kazal proteinase inhibitor homology <KPI9>
19-992/Region: serine/threonine-rich
184-1086/Region: motor neuron attachment (L-R-E) motif
.47-1215/Region: serine/threonine-rich
24-1257/Domain: EGF homology <EG1>
287-1442/Domain: laminin G repeat homology <LG1>
144-1476/Domain: EGF homology <EG2>
183-1515/Domain: EGF homology <EG3>
55-1706/Domain: laminin G repeat homology <LG2>
113-1747/Domain: EGF homology <EG4>
107-1959/Domain: laminin G repeat homology <LG3>
1-116,105-137,171-191,180-212,244-263,252-284,316-335,324-356,389-408,397-429,454-473
16,1483-1494,1488-1504,1506-1515/Disulfide bonds: #status predicted
15,672,827,957/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.0%; Score 38; DB 1; Length 1959;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

3 CXXIYXPVC 11
| | | | |
389 CSGIYDPVC 397

JLT 11
138
nucoid, third domain - Coqui francolin (fragment)
Species: Francolinus coqui (Coqui francolin)
Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 07-Feb-1997
Accession: G31438
Askowski Jr., M.; Kato, I.; Ardelt, W.; Cook, J.; Denton, A.; Empie, M.W.; Kohr, W.J.;
; Wiczorek, M.
Chemistry 26, 202-221, 1987
Title: Ovomuroid third domains from 100 avian species: isolation, sequences, and hyper
Reference number: A90515; MUID:87157615; PMID:3828298
Accession: G31438
Molecule type: protein
Residues: 1-53 <LAS>
Note: the authors designate this sequence with the code OMCOF3
Superfamily: ovomucoid; Kazal proteinase inhibitor homology
Keywords: egg white; glycoprotein; serine proteinase inhibitor
53/Domain: Kazal proteinase inhibitor homology <KPI>
35,13-32,21-53/Disulfide bonds: #status predicted
5/Binding site: carbohydrate (Asn) (covalent) #status absent
2/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental

Query Match 74.0%; Score 37; DB 2; Length 53;
Best Local Similarity 60.0%; Pred. No. 1.8;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

2 GCXXIYXPVC 11
| | | | |
12 GCTMEYRPVC 21

ULT 12
340
tate-gated sodium channel homolog - Bdeloura candida
Species: Bdeloura candida
Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
Accession: T31340
eziorski, M.C.; Greenberg, R.M.; Anderson, P.A.

submitted to the EMBL Data Library, March 1997
A;Description: A putative voltage-gated sodium channel from the turbellarian flatworm Bde
A;Reference number: Z21006
A;Accession: T31340
A;Status: preliminary; translated from GB/EMBL/DBDJB
A;Molecule type: mRNA
A;Residues: 1-1699 <JEZ>
A;Cross-references: EMBL:U93074; NID:G1947093; PID:G1947094; PIDN:AAC63049.1
C;Genetics:
A;Gene: Nal
C;Superfamily: sodium channel protein

Query Match 72.0%; Score 36; DB 2; Length 1699;
Best Local Similarity 55.6%; Pred. No. 50;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11
| | | | |
Db 469 CDDIYQPIC 477

RESULT 13
TIRT2
pancreatic secretory trypsin inhibitor II precursor - rat
N;Alternate names: hepatic proteinase inhibitor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jun-2000
C;Accession: B33292; S28947; S00634; S08982
R;Horii, A.; Tomita, N.; Yokouchi, H.; Doi, S.; Uda, K.; Ogawa, M.; Mori, T.; Matsubara,
Biochem. Biophys. Res. Commun. 162, 151-159, 1989
A;Title: On the cDNA's for two types of rat pancreatic secretory trypsin inhibitor.
A;Reference number: A33292; MUID:89322236; PMID:2751646
A;Accession: B33292
A;Molecule type: mRNA
A;Residues: 1-79 <HOR>
A;Cross-references: GB:M27883
R;Tsuzuki, S.; Miura, Y.; Fushiki, T.; Oomori, T.; Satoh, T.; Natori, Y.; Sugimoto, E.
Biochim. Biophys. Acta 1132, 199-202, 1992
A;Title: Molecular cloning and characterization of genes encoding rat pancreatic cholecys
A;Reference number: S28946; MUID:93003324; PMID:1390891
A;Accession: S28947
A;Molecule type: DNA
A;Residues: 1-79 <TSU>
A;Cross-references: DDBJ:D11325; NID:G220885; PIDN:BAA01945.1; PID:G220887
R;Uda, K.I.; Ogawa, M.; Shibata, T.; Murata, A.; Mori, T.; Kikuchi, N.; Yoshida, N.; Tsur
Biol. Chem. Hoppe-Seyler 369(Suppl.), 55-61, 1988
A;Title: Purification, characterization and amino-acid sequencing of two pancreatic secr
A;Reference number: S00633; MUID:89076534; PMID:3202973
A;Accession: S00634
A;Molecule type: protein
A;Residues: 24-79 <UDA>
R;Kido, H.; Yokogoshi, Y.; Katunuma, N.
Eur. J. Biochem. 188, 501-506, 1990
A;Title: A low-molecular-mass Kazal-type protease inhibitor isolated from rat hepatocyte
A;Reference number: S08982; MUID:90235819; PMID:2110056
A;Accession: S08982
A;Molecule type: protein
A;Residues: 24-79 <KID>
C;Genetics:
A;Introns: 19/1; 29/3; 65/2
C;Superfamily: pancreatic secretory trypsin inhibitor; Kazal proteinase inhibitor homolo
C;Keywords: pancreas; serine proteinase inhibitor
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-79/Product: pancreatic secretory trypsin inhibitor I #status experimental <MAT>
F;30-79/Domain: Kazal proteinase inhibitor homology <KPI>
F;32-61,39-58,47-79/Disulfide bonds: #status predicted
F;41/Inhibitory site: Arg (trypsin) #status predicted

Query Match 70.0%; Score 35; DB 1; Length 79;
Best Local Similarity 60.0%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11
```

Db 38 GCPDYDPVC 47  
|| | |||

## RESULT 14

JA0128  
hypothetical 11K protein - potato virus S  
A;Species: potato virus S  
C;Note: host Chenopodium quinoa  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Jun-2000  
C;Accession: JA0128  
R;MacKenzie, D.J.; Tremaine, J.H.; Stace-Smith, R.  
J. Gen. Virol. 70, 1053-1063, 1989  
A;Title: Organization and intervirial homologies of the 3'-terminal portion of potato virus S  
A;Reference number: JA0123; MUID:89279283; PMID:2732711  
A;Accession: JA0128  
A;Molecule type: genomic RNA  
A;Residues: 1-93 <MAC>  
A;Cross-references: GB:D00461; NID:G222438; PIDN:BAA00356.1; PID:G2160373  
C;Comment: The genome is a single-stranded, positive-sense RNA.  
C;Superfamily: potato virus nucleic acid-binding protein  
C;Keywords: DNA binding; zinc finger

Query Match 70.0%; Score 35; DB 2; Length 93;  
Best Local Similarity 55.6%; Pred. No. 7;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
| : |||  
Db 57 CYRVYPPVC 65

## RESULT 15

T18525  
diacylglycerol kinase (EC 2.7.1.107) eta - hamster  
N;Alternate names: diglyceride kinase  
C;Species: Cricetinae gen. sp. (hamster)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
C;Accession: T18525  
R;Klauck, T.M.; Xu, X.; Mousseau, B.; Jaken, S.  
J. Biol. Chem. 271, 19781-19788, 1996  
A;Title: Cloning and characterization of a glucocorticoid-induced diacylglycerol kinase.  
A;Reference number: Z18948; MUID:96355275; PMID:8702685  
A;Accession: T18525  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1154 <KLA>  
A;Cross-references: EMBL:U59429; NID:G1401231; PID:G1401232; PIDN:AAC52714.1  
A;Experimental source: strain Syrian  
C;Superfamily: protein kinase C zinc-binding repeat homology  
C;Keywords: phosphotransferase  
F;170-219/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 70.0%; Score 35; DB 2; Length 1154;  
Best Local Similarity 55.6%; Pred. No. 57;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
| : |||  
Db 284 CKDLYHPVC 292

Search completed: August 11, 2004, 13:09:59  
Job time : 13.0476 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 11, 2004, 13:01:12 ; Search time 8.38095 Seconds  
(without alignments)  
68.342 Million cell updates/sec

le: US-10-661-207-2  
fect score: 50  
ence: 1 XGXXIYXPVC 11

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 141681 seqs, 52070155 residues

al number of hits satisfying chosen parameters: 141681

imum DB seq length: 0  
imum DB seq length: 2000000000

--processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

abase : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

alt No.	Score	Query Match	Length	ID	Description
1	44	88.0	56	1 IPK1 HORSE	P81634 equus cabal
2	44	88.0	79	1 IPK1 HUMAN	P00995 homo sapien
3	42	84.0	56	1 IPK1 BOVIN	P00996 bos taurus
4	42	84.0	56	1 IPK1 PIG	P00998 sus scrofa
5	42	84.0	56	1 IPK1 SHEEP	P00997 ovis aries
6	42	84.0	57	1 IPK1 CANFA	P04542 canis famil
7	42	84.0	79	1 IPK1 RAT	P09655 rattus norv
8	42	84.0	80	1 ISK3 MOUSE	P09036 mus musculu
9	40	80.0	69	1 IPK1 STRCA	Q9psm2 struthio ca
10	39	78.0	197	1 MCP1 MELCP	P82968 melithaea c
11	38	76.0	86	1 ISK4 PIG	P37109 sus scrofa
12	38	76.0	156	1 ESM1 DROME	O97176 drosophila
13	38	76.0	472	1 IOV7 CHICK	P10184 gallus gall
14	38	76.0	1959	1 AGRI RAT	P25304 rattus norv
15	37	74.0	53	1 IOVO FRACO	P05595 francolinus
16	36	72.0	351	1 DPGN DIPMA	O96790 dipetalogas
17	36	72.0	6684	1 RIAB CVPPU	Q9iw06 p replicase
18	36	72.0	6781	1 RIAB PEDV7	Q9lav2 p replicase
19	35	70.0	79	1 IPK2 RAT	P09656 rattus norv
20	35	70.0	93	1 VNBV PVSP	P16654 potato viru
21	35	70.0	140	1 VNBV LSV	P27336 lily sympto
22	35	70.0	1154	1 KDGD MESAU	Q64398 mesocricetu
23	33	66.0	54	1 IOVO GEOCA	P05615 geococcyx c
24	32	64.0	223	1 KITH UREPA	Q9ppp5 ureaplasma
25	32	64.0	806	1 BIMA EMENI	P17885 emericelella
26	32	64.0	1955	1 AGRI CHICK	P31696 gallus gall
27	31.5	63.0	759	1 PMT6 YEAST	P42934 saccharomyc
28	31	62.0	48	1 IELA ANESU	P16895 anemonia su
29	31	62.0	53	1 IOVO ARBTO	P05601 arborophila
30	31	62.0	53	1 IOVO FRAAF	P05594 francolinus
31	31	62.0	56	1 IOVO FRAER	P05596 francolinus
32	31	62.0	81	1 IAC MACFA	P34953 macaca fasc
33	31	62.0	84	1 IAC2 HUMAN	P20155 homo sapien

ALIGNMENTS

RESULT 1

IPK1\_HORSE  
ID IPK1\_HORSE STANDARD; PRT; 56 AA.  
AC P81634;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN SPINK1 OR PSTI.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Pancreas;  
RA Voigt J.;  
RT "Representation and characterization of pancreatic secretory  
proteinase inhibitors from various animal species.";  
RL Thesis (1993), Ludwig-Maximilians University / Munich, Germany.  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
DR HSSP; P37109; IPCE.  
DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor.  
FT DOMAIN 7 56 KAZAL-LIKE.  
FT DISULFID 9 38 BY SIMILARITY.  
FT DISULFID 16 35 BY SIMILARITY.  
FT DISULFID 24 56 BY SIMILARITY.  
FT ACT\_SITE 18 19 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 56 AA; 6093 MW; D0DAA7E4FDB507E0 CRC64;

Query Match 88.0%; Score 44; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.014;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
Db 15 GCTKIYNPVC 24

RESULT 2

IPK1\_HUMAN  
ID IPK1\_HUMAN STANDARD; PRT; 79 AA.  
AC P00995;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)



DE Pancreatic secretory trypsin inhibitor precursor (Tumor-associated  
DE trypsin inhibitor) (TATI) (Serine protease inhibitor Kazal-type 1).  
GN SPINK1 OR PSTI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88106485; PubMed=3501289;  
RA Horii A., Kobayashi T., Tomita N., Yamamoto T., Fukushima S.,  
RA Murotsu T., Ogawa M., Mori T., Matsubara K.;  
RT "Primary structure of human pancreatic secretory trypsin inhibitor  
RT (PSTI) gene.";  
RL Biochem. Biophys. Res. Commun. 149:635-641(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86050645; PubMed=3877508;  
RA Yamamoto T., Nakamura Y., Nishide T., Emi M., Ogawa M., Mori T.,  
RA Matsubara K.;  
RT "Molecular cloning and nucleotide sequence of human pancreatic  
RT secretory trypsin inhibitor (PSTI) cDNA.";  
RL Biochem. Biophys. Res. Commun. 132:605-612(1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88083571; PubMed=2961612;  
RA Tomita N., Horii A., Yamamoto T., Ogawa M., Mori T., Matsubara K.;  
RT "Expression of pancreatic secretory trypsin inhibitor gene in  
RT neoplastic tissues.";  
RL FEBS Lett. 225:113-119(1987).  
RN [4]  
RP SEQUENCE FROM N.A., VARIANTS HPC PRO-14 AND SER-34, AND VARIANT  
RP SER-55.  
RX MEDLINE=20296628; PubMed=10835640;  
RA Witt H., Luck W., Hennies H.C., Classen M., Kage A., Lass U.,  
RA Landt O., Becker M.;  
RT "Mutations in the gene encoding the serine protease inhibitor, Kazal  
RT type 1 are associated with chronic pancreatitis.";  
RL Nat. Genet. 25:213-216(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas, and Spleen;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP SEQUENCE OF 24-79.  
RX MEDLINE=77133145; PubMed=843082;  
RA Bartelt D.C., Shapanka R., Greene L.J.;  
RT "The primary structure of the human pancreatic secretory trypsin  
RT inhibitor. Amino acid sequence of the reduced S-aminoethylated  
RT protein.";  
RL Arch. Biochem. Biophys. 179:189-199(1977).  
RN [7]  
RP SEQUENCE OF 24-46.

RX MEDLINE=83056875; PubMed=7142173;  
RA Huhtala M.-L., Pesonen K., Kalkkinen N., Stenman U.-H.;  
RT "Purification and characterization of a tumor-associated trypsin  
RT inhibitor from the urine of a patient with ovarian cancer.";  
RL J. Biol. Chem. 257:13713-13716(1982).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RX MEDLINE=92309406; PubMed=1613792;  
RA Hecht H.-J., Szardenings M., Collins J., Schomburg D.;  
RT "Three-dimensional structure of a recombinant variant of human  
RT pancreatic secretory trypsin inhibitor (Kazal type).";  
RL J. Mol. Biol. 225:1095-1103(1992).  
RN [9]  
RP STRUCTURE BY NMR OF MUTANT LEU-41/ARG-44.  
RX MEDLINE=93164251; PubMed=8433367;  
RA Klaus W., Schomburg D.;  
RT "Solution structure of a variant of human pancreatic secretory  
RT trypsin inhibitor determined by nuclear magnetic resonance  
RT spectroscopy.";  
RL J. Mol. Biol. 229:695-706(1993).  
RN [10]  
RP VARIANT HPC SER-34, AND VARIANT SER-55.  
RX MEDLINE=20144835; PubMed=10691414;  
RA Chen J.-M., Mercier B., Audrezet M.-P., Ferec C.;  
RT "Mutational analysis of the human pancreatic secretory trypsin  
RT inhibitor (PSTI) gene in hereditary and sporadic chronic  
RT pancreatitis.";  
RL J. Med. Genet. 37:67-69(2000).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
CC is to prevent the trypsin-catalyzed premature activation of  
CC zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DISEASE: Defects in SPINK1 are a cause of hereditary pancreatitis,  
CC (HPC); also known as chronic pancreatitis (CP). HPC is an  
CC autosomal dominant disease characterized by the presence of  
CC calculi in pancreatic ducts. It causes severe abdominal pain  
CC attacks.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M20530; AAA36522.1; -.  
CC EMBL; M22971; AAA36522.1; JOINED.  
CC EMBL; M20528; AAA36522.1; JOINED.  
CC EMBL; M20529; AAA36522.1; JOINED.  
CC EMBL; Y00705; CAA68697.1; -.  
CC EMBL; M11949; AAA36521.1; -.  
CC EMBL; AF286028; AAG00531.1; -.  
CC EMBL; BC025790; AAH25790.1; -.  
CC PIR; A27484; TIHUA.  
CC PDB; 1HPT; 31-OCT-93.  
CC Genew; HGNC:11244; SPINK1.  
CC MIM; 167790; -.  
CC MIM; 167800; -.  
CC GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.  
CC InterPro; IPR002350; kazal.  
CC InterPro; IPR001239; Kazal\_inhib.  
CC Pfam; PF00050; kazal; 1.  
CC PRINTS; PR00290; KAZALINHBTR.  
CC SMART; SM00280; KAZAL; 1.  
CC PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor; Signal; Disease mutation; Polymorphism;  
3D-structure.  
FT SIGNAL 1 23  
FT CHAIN 24 79 PANCREATIC SECRETORY TRYPSIN INHIBITOR.  
FT DOMAIN 30 79 KAZAL-LIKE.  
FT DISULFID 32 61



DISULFID 39 58  
DISULFID 47 79  
ACT\_SITE 41 42  
VARIANT 14 14  
REACTIVE BOND.  
L -> P (IN HPC).  
/FTID=VAR\_011688.  
N -> S (IN HPC).  
/FTID=VAR\_011689.  
P -> S.  
/FTID=VAR\_011690.  
D -> N (IN REF. 6 AND 7).  
N -> D (IN REF. 6).  
N -> G (IN REF. 3).  
SEQUENCE 79 AA; 8507 MW; 3583C8196952EB3A CRC64;  
very Match 88.0%; Score 44; DB 1; Length 79;  
est Local Similarity 70.0%; Pred. No. 0.02;  
atches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
2 GCXXIYXPVC 11  
|||  
38 GCTKIYDPVC 47

ULT 3  
1\_BOVIN STANDARD; PRT; 56 AA.  
P00996;  
21-JUL-1986 (Rel. 01, Created)  
21-JUL-1986 (Rel. 01, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Pancreatic secretory trypsin inhibitor.  
SPINK1 OR PSTI.  
Bos taurus (Bovine).  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
NCBI\_TaxID=9913;  
[1]  
SEQUENCE.  
MEDLINE=69187206; PubMed=5769997;  
Greene L.J., Bartelt D.C.;  
"The structure of the bovine pancreatic secretory trypsin inhibitor --  
Kazal's inhibitor. II. The order of the tryptic peptides.";  
J. Biol. Chem. 244:2646-2657(1969).  
[2]  
DISULFIDE BONDS.  
MEDLINE=72086018; PubMed=5135319;  
Guy O., Shapanka R., Greene L.J.;  
"The structure of the bovine pancreatic secretory trypsin inhibitor --  
Kazal's inhibitor. 3. Determination of the disulfide bonds and  
proteolysis by thermolysin.";  
J. Biol. Chem. 246:7740-7747(1971).  
-!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- SIMILARITY: Contains 1 Kazal-like domain.  
PIR; A01230; TIBOA.  
HSSP; P37109; LPCE.  
InterPro; IPR002350; kazal.  
InterPro; IPR001239; Kazal\_inhib.  
Pfam; PF00050; Kazal; 1.  
PRINTS; PR00290; KAZALINHBTR.  
SMART; SM00280; KAZAL; 1.  
PROSITE; PS00282; KAZAL; 1.  
Serine protease inhibitor.  
DOMAIN 7 56 KAZAL-LIKE.  
DISULFID 9 38  
DISULFID 16 35  
DISULFID 24 56  
ACT SITE 18 19 REACTIVE BOND.  
SEQUENCE 56 AA; 6161 MW; AB64A512AC851B26 CRC64;

Query Match 84.0%; Score 42; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.036;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GCXXIYXPVC 11  
|||  
Db 15 GCPRIYNPVC 24

RESULT 4  
IPK1\_PIG STANDARD; PRT; 56 AA.  
AC P00998;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN SPINK1 OR PSTI.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=70283430; PubMed=5466061;  
Tschesche H., Wächter E.;  
"The structure of the porcine pancreatic secretory trypsin inhibitor.  
I. A sequence determination by Edman degradation and mass spectral  
identification of the p-bromophenyl-thiohydantoins.";  
Eur. J. Biochem. 16:187-198(1970).  
[2]  
RP SEQUENCE.  
RX MEDLINE=71178430; PubMed=5103069;  
Bartelt D.C., Greene L.J.;  
"The primary structure of the porcine pancreatic secretory trypsin  
inhibitor. I. Amino acid sequence of the reduced S-aminoethylated  
protein.";  
J. Biol. Chem. 246:2218-2229(1971).  
[3]  
RP DISULFIDE BONDS.  
RX MEDLINE=73001409; PubMed=4672150;  
Tschesche H., Schneider M., Reidel G., Klein H.;  
"Disulfide bridges of the secretory trypsin inhibitor from porcine  
pancreas and the degradation of covalent structure during the  
temporary inhibition.";  
Hoppe-Seyler's Z. Physiol. Chem. 353:763-764(1972).  
[4]  
RP SEQUENCE OF A SECOND INHIBITOR.  
RX MEDLINE=71092915; PubMed=5531651;  
Tschesche H., Wächter E.;  
"Trypsin inhibitors. VII. Primary structure of the specific trypsin  
inhibitor II (Kazal-type) from porcine pancreas. Sequence analysis  
with mass spectrometry identification of the  
p-bromophenylthio-hydantoins from the Edman degradation.";  
Hoppe-Seyler's Z. Physiol. Chem. 351:1449-1459(1970).  
[5]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
RX MEDLINE=83189075; PubMed=7169635;  
Bolognesi M., Gatti G., Menegatti E., Guarneri M., Marquart M.,  
Papamokos E., Huber R.;  
"Three-dimensional structure of the complex between pancreatic  
secretory trypsin inhibitor (Kazal type) and trypsinogen at 1.8-A  
resolution. Structure solution, crystallographic refinement and  
preliminary structural interpretation.";  
J. Mol. Biol. 162:839-868(1982).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
DR PIR; A91174; TPG.  
DR PDB; 1TGS; 14-MAR-85.  
DR InterPro; IPR002350; kazal.

DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor; 3D-structure.  
FT DOMAIN 7 56 KAZAL-LIKE.  
FT DISULFID 9 38  
FT DISULFID 16 35  
FT DISULFID 24 56  
FT ACT\_SITE 18 19  
FT VARIANT 1 4  
FT STRAND 16 17  
FT STRAND 23 25  
FT TURN 26 27  
FT STRAND 30 31  
FT HELIX 34 42  
FT TURN 43 43  
FT STRAND 50 53  
SQ SEQUENCE 56 AA; 6023 MW; 39A3649DADF16D25 CRC64;  
  
Query Match 84.0%; Score 42; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.036;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 GCXXIYXPVC 11  
Db 15 GCPKIYNPVC 24  
  
RESULT 5  
IPK1\_SHEEP STANDARD; PRT; 56 AA.  
AC P00997;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN SPINK1 OR PSTI.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP PARTIAL SEQUENCE.  
RA Teschesche H., Obermeier R., Hochstrasser K.;  
RL Unpublished results, cited by:  
RL Teschesche H., Wachter E., Kupfer S., Obermeier R., Reidel G.,  
RL Haenisch G., Schneider M.;  
RL (In) Fritz H., Tschesche H. (eds.);  
RL Proceedings of the international research conference on proteinase  
inhibitors, pp.207-222, Walter de Gruyter, New York (1971).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
DR PIR; A94487; TISHA.  
DR HSSP; P37109; 1PCE.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor.  
FT DOMAIN 7 56 KAZAL-LIKE.  
FT DISULFID 9 38 BY SIMILARITY.  
FT DISULFID 16 35 BY SIMILARITY.  
FT DISULFID 24 56 BY SIMILARITY.  
FT ACT\_SITE 18 19 REACTIVE BOND.  
SQ SEQUENCE 56 AA; 6145 MW; B8459712AC851B26 CRC64;

DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor; 3D-structure.  
FT DOMAIN 7 56 KAZAL-LIKE.  
FT DISULFID 9 38  
FT DISULFID 16 35  
FT DISULFID 24 56  
FT ACT\_SITE 18 19  
FT VARIANT 1 4  
FT STRAND 16 17  
FT STRAND 23 25  
FT TURN 26 27  
FT STRAND 30 31  
FT HELIX 34 42  
FT TURN 43 43  
FT STRAND 50 53  
SQ SEQUENCE 56 AA; 6023 MW; 39A3649DADF16D25 CRC64;  
  
Query Match 84.0%; Score 42; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.036;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 GCXXIYXPVC 11  
Db 15 GCPKIYNPVC 24  
  
RESULT 5  
IPK1\_SHEEP STANDARD; PRT; 56 AA.  
AC P00997;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN SPINK1 OR PSTI.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP PARTIAL SEQUENCE.  
RA Teschesche H., Obermeier R., Hochstrasser K.;  
RL Unpublished results, cited by:  
RL Teschesche H., Wachter E., Kupfer S., Obermeier R., Reidel G.,  
RL Haenisch G., Schneider M.;  
RL (In) Fritz H., Tschesche H. (eds.);  
RL Proceedings of the international research conference on proteinase  
inhibitors, pp.207-222, Walter de Gruyter, New York (1971).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
DR PIR; A94487; TISHA.  
DR HSSP; P37109; 1PCE.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor.  
FT DOMAIN 7 56 KAZAL-LIKE.  
FT DISULFID 9 38 BY SIMILARITY.  
FT DISULFID 16 35 BY SIMILARITY.  
FT DISULFID 24 56 BY SIMILARITY.  
FT ACT\_SITE 18 19 REACTIVE BOND.  
SQ SEQUENCE 56 AA; 6145 MW; B8459712AC851B26 CRC64;

Query Match 84.0%; Score 42; DB 1; Length 56;  
Best Local Similarity 70.0%; Pred. No. 0.036;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 GCXXIYXPVC 11  
Db 15 GCPRIYNPVC 24  
  
RESULT 6  
IPK1\_CANFA STANDARD; PRT; 57 AA.  
AC P04542;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN SPINK1 OR PSTI.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=86030679; PubMed=4054311;  
RA Kikuchi N., Nagata K., Yoshida N., Tanaka T., Yamamoto M., Saitoh Y.;  
RT "Purification and complete amino acid sequence of canine pancreatic  
secretory trypsin inhibitor.";  
RL FEBS Lett. 191:269-272(1985).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MISCELLANEOUS: This protein is unusual, with respect to the other  
mammalian PSTI proteins, in having an extra N-terminal residue.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
DR PIR; A01232; TIDGA.  
DR HSSP; P37109; 1PCE.  
DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor.  
FT DOMAIN 8 57 KAZAL-LIKE.  
FT DISULFID 10 39 BY SIMILARITY.  
FT DISULFID 17 36 BY SIMILARITY.  
FT DISULFID 25 57 BY SIMILARITY.  
FT ACT\_SITE 19 20 REACTIVE BOND.  
SQ SEQUENCE 57 AA; 6319 MW; 2DB674A2FD98CD43 CRC64;  
  
Query Match 84.0%; Score 42; DB 1; Length 57;  
Best Local Similarity 60.0%; Pred. No. 0.037;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 GCXXIYXPVC 11  
Db 16 GCNKIYNPIC 25  
  
RESULT 7  
IPK1\_RAT STANDARD; PRT; 79 AA.  
AC P09655; P13072;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor 1 precursor (PSTI-1)  
DE (Cholecystokinin-releasing peptide) (Monitor peptide).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas;  
MEDLINE=90098786; PubMed=2602119;  
Fukuoka S.-I., Scheele G.A.;  
"Complementary nucleotide sequence for monitor peptide, a novel  
cholecystokinin-releasing peptide in the rat.";  
Nucleic Acids Res. 17:10111-10111(1989).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=Wistar; TISSUE=Pancreas;  
MEDLINE=90093122; PubMed=2293709;  
Fukuoka S.-I., Scheele G.A.;  
"Rapid and selective cloning of monitor peptide, a novel  
cholecystokinin-releasing peptide, using minimal amino acid sequence  
and the polymerase chain reaction.";  
Pancreas 5:1-7(1990).  
[3]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas;  
MEDLINE=89322236; PubMed=2751646;  
Hori A., Tomita N., Yokouchi H., Doi S., Uda K., Ogawa M., Mori T.,  
Matsubara K.;  
"On the cDNA's for two types of rat pancreatic secretory trypsin  
inhibitor.";  
Biochem. Biophys. Res. Commun. 162:151-159(1989).  
[4]  
SEQUENCE FROM N.A.  
MEDLINE=91293130; PubMed=2065678;  
Tsuzuki S., Fushiki T., Kondo A., Murayama H., Sugimoto E.;  
"Effect of a high-protein diet on the gene expression of a trypsin-  
sensitive, cholecystokinin-releasing peptide (monitor peptide) in the  
pancreas.";  
Eur. J. Biochem. 199:245-252(1991).  
[5]  
SEQUENCE FROM N.A.  
MEDLINE=93003324; PubMed=1390891;  
Tsuzuki S., Miura Y., Fushiki T., Omori T., Satch T., Natori Y.,  
Sugimoto E.;  
"Molecular cloning and characterization of genes encoding rat  
pancreatic cholecystokinin (CCK)-releasing peptide (monitor peptide)  
and pancreatic secretory trypsin inhibitor (PSTI).";  
Biochim. Biophys. Acta 1132:199-202(1992).  
[6]  
SEQUENCE OF 19-79.  
STRAIN=Wistar; TISSUE=Pancreas;  
MEDLINE=89076534; PubMed=3202973;  
Uda K., Ogawa M., Shibata T., Murata A., Mori T., Kikuchi N.,  
Yoshida N., Tsunawasa S., Sakiyama F.;  
"Purification, characterization and amino-acid sequencing of two  
pancreatic secretory trypsin inhibitors in rat pancreatic juice.";  
Biol. Chem. Hoppe-Seyler 369:55-61(1988).  
[7]  
SEQUENCE OF 19-79.  
TISSUE=Pancreas;  
MEDLINE=87250528; PubMed=3597401;  
Iwai K., Fukuoka S.-I., Fushiki T., Tsujikawa M., Hirose M.,  
Tsunasawa S., Sakiyama F.;  
"Purification and sequencing of a trypsin-sensitive cholecystokinin-  
releasing peptide from rat pancreatic juice. Its homology with  
pancreatic secretory trypsin inhibitor.";  
J. Biol. Chem. 262:8956-8959(1987).  
-!- FUNCTION: This is a trypsin inhibitor, its physiological function  
is to prevent the trypsin-catalyzed premature activation of  
zymogens within the pancreas.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- SIMILARITY: Contains 1 Kazal-like domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X59696; CAA42217.1; -  
DR EMBL; M22162; AAA41629.1; -  
DR EMBL; M35299; AAA74479.1; -  
DR EMBL; M35300; AAA41977.1; -  
DR EMBL; M27882; AAA41975.1; -  
DR EMBL; D11321; BAA01944.1; -  
DR PIR; S09602; TIRT1.  
DR HSSP; P37109; IPCE.  
DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor; Signal.  
FT SIGNAL 1 18 PANCREATIC SECRETORY TRYPSIN INHIBITOR I.  
FT CHAIN 19 79 KAZAL-LIKE  
FT DOMAIN 30 79 BY SIMILARITY.  
FT DISULFID 32 61 BY SIMILARITY.  
FT DISULFID 39 58 BY SIMILARITY.  
FT DISULFID 47 79 REACTIVE BOND.  
FT ACT\_SITE 41 42 T -> G (IN REF. 7).  
FT CONFLICT 78 78 5816D55DF7B57874 CRC64;  
SQ SEQUENCE 79 AA; 8528 MW; 5816D55DF7B57874 CRC64;  
  
Query Match 84.0%; Score 42; DB 1; Length 79;  
Best Local Similarity 70.0%; Pred. No. 0.05;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 GCXXIYXPVC 11  
||| |||||  
Db 38 GCPRIYDPVC 47  
  
RESULT B  
ISK3\_MOUSE  
ID ISK3\_MOUSE STANDARD; PRT; 80 AA.  
AC P09036;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serine protease inhibitor Kazal-type 3 precursor (Prostatic secretory  
glycoprotein) (P12).  
DE SPINK3.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE=88111560; PubMed=3428272;  
RA Mills J.S., Needham M., Parker M.G.;  
RT "A secretory protease inhibitor requires androgens for its expression  
in male sex accessory tissues but is expressed constitutively in  
pancreas.";  
RL EMBO J. 6:3711-3717(1987).  
CC -!- FUNCTION: SERINE PROTEASE INHIBITOR WHICH EXHIBITS ANTI-TRYPSIN  
ACTIVITY.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- INDUCTION: By androgens.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X06342; CAA29648.1; --  
DR PIR; S01498; S01498.  
DR HSSP; P37109; 1PCE.

DR MGD; MGI:106202; Spink3.  
DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.

DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.

DR Serine protease inhibitor; Glycoprotein; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 80 SERINE PROTEASE INHIBITOR KAZAL-TYPE 3.

FT DOMAIN 31 80 KAZAL-LIKE.  
FT DISULFID 33 62 BY SIMILARITY.  
FT DISULFID 40 59 BY SIMILARITY.

FT DISULFID 48 80 BY SIMILARITY.  
FT ACT\_SITE 42 43 REACTIVE BOND.  
SQ SEQUENCE - 80 AA; 8488 MW; 4DC1F2EC4804CCA6 CRC64;

Query Match 84.0%; Score 42; DB 1; Length 80;  
Best Local Similarity 70.0%; Pred. No. 0.051;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
Db 39 GCPRIYDPVC 48

RESULT 9  
IPK1\_STRCA STANDARD; PRT; 69 AA.

AC Q9PSM2;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN PSTI.

OS Struthio camelus (Ostrich).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;

OC Struthio.  
OX NCBI\_TaxID=8801;  
RN [1]

RP SEQUENCE.  
RC TISSUE=Pancreas;  
RX MEDLINE=97026357; PubMed=8872536;

RA Zhao M., Naude R.J., Muramoto K., Oelofsen W.;  
RT "Purification and characterization of ostrich pancreatic secretory  
RT trypsin inhibitor."

RL Int. J. Pept. Protein Res. 48:174-181(1996).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
CC is to prevent the trypsin-catalyzed premature activation of

CC zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.

DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.

DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.

DR Serine protease inhibitor.  
KW DOMAIN 12 63 KAZAL-LIKE.  
FT DISULFID 14 45 BY SIMILARITY.

FT DISULFID 23 42 BY SIMILARITY.  
FT DISULFID 31 63 BY SIMILARITY.  
FT ACT\_SITE 25 26 REACTIVE BOND.

SQ SEQUENCE 69 AA; 7656 MW; CFB8AE73BFC148870 CRC64;

Query Match 84.0%; Score 42; DB 1; Length 80;  
Best Local Similarity 70.0%; Pred. No. 0.051;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
Db 39 GCPRIYDPVC 48

RESULT 9  
IPK1\_STRCA STANDARD; PRT; 69 AA.

AC Q9PSM2;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pancreatic secretory trypsin inhibitor.  
GN PSTI.

OS Struthio camelus (Ostrich).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;

OC Struthio.  
OX NCBI\_TaxID=8801;  
RN [1]

RP SEQUENCE.  
RC TISSUE=Pancreas;  
RX MEDLINE=97026357; PubMed=8872536;

RA Zhao M., Naude R.J., Muramoto K., Oelofsen W.;  
RT "Purification and characterization of ostrich pancreatic secretory  
RT trypsin inhibitor."

RL Int. J. Pept. Protein Res. 48:174-181(1996).  
CC -!- FUNCTION: This is a trypsin inhibitor, its physiological function  
CC is to prevent the trypsin-catalyzed premature activation of

CC zymogens within the pancreas.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.

DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.

DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.

DR Serine protease inhibitor.  
KW DOMAIN 12 63 KAZAL-LIKE.  
FT DISULFID 14 45 BY SIMILARITY.

Query Match 80.0%; Score 40; DB 1; Length 69;  
Best Local Similarity 60.0%; Pred. No. 0.11;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
Db 22 GCAKIFDPVC 31

RESULT 10  
MCPI\_MELCP STANDARD; PRT; 197 AA.

AC P82968;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protease inhibitor (McaPI).  
OS Melithaea caledonica.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Gorgonacea;  
OC Scleraxonia; Melithaeidae; Melithaea.  
OX NCBI\_TaxID=156534;

RN [1]  
RP SEQUENCE, AND VARIANT VAL-195.  
RA Peduzzi J., Longeon A., Guyot M., Barthelemy M.;

RT "Amino acid sequence and kinetic properties of a four-domain proteases  
RT inhibitor from the gorgonian Melithaea caledonica.";  
RL Eur. J. Biochem. 0:0-0(2001).

CC -!- FUNCTION: INHIBITS TRYPSIN, KALLIKREIN, SUBTILISIN CARLSBERG,  
CC HUMAN LEUKOCYTE ELASTASE, PORCINE PANCREATIC ELASTASE AND  
CC CHYMOTRYPSIN. TWO DOMAINS ARE FOR THE INHIBITION OF CHYMOTRYPSIN.

CC -!- SIMILARITY: Contains 3 Kazal-like domains.  
CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
DR HSSP; P31713; 1SHP.

DR InterPro; IPR002350; kazal.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR Pfam; PF00050; kazal; 3.

DR Pfam; PF00014; Kunitz\_BPTI; 1.  
DR PRINTS; PR00759; BASICTPTASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 1.

DR SMART; SM00280; KAZAL; 3.  
DR SMART; SM00131; KU; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

DR PROSITE; PS00279; BPTI\_KUNITZ\_2; 1.  
DR PROSITE; PS00282; KAZAL; FALSE\_NEG.  
KW Serine protease inhibitor; Repeat.

FT DOMAIN 1 45 KAZAL-LIKE 1.  
FT DOMAIN 49 93 KAZAL-LIKE 2.  
FT DOMAIN 97 141 KAZAL-LIKE 3.

FT DOMAIN 144 194 BPTI/KUNITZ INHIBITOR.  
FT ACT\_SITE 7 8 REACTIVE BOND 1 (BY SIMILARITY).  
FT ACT\_SITE 55 56 REACTIVE BOND 2 (BY SIMILARITY).

FT ACT\_SITE 103 104 REACTIVE BOND 3 (BY SIMILARITY).  
FT ACT\_SITE 154 155 REACTIVE BOND 4 (BY SIMILARITY).  
FT DISULFID 1 31 BY SIMILARITY.

FT DISULFID 5 24 BY SIMILARITY.  
FT DISULFID 13 45 BY SIMILARITY.  
FT DISULFID 49 79 BY SIMILARITY.

FT DISULFID 53 72 BY SIMILARITY.  
FT DISULFID 61 93 BY SIMILARITY.  
FT DISULFID 97 127 BY SIMILARITY.

FT DISULFID 101 120 BY SIMILARITY.  
FT DISULFID 109 141 BY SIMILARITY.  
FT DISULFID 144 194 BY SIMILARITY.

FT DISULFID 153 177 BY SIMILARITY.  
FT DISULFID 169 190 BY SIMILARITY.  
FT VARIANT 195 195 M -> V.

SQ SEQUENCE 197 AA; 21248 MW; 199D08A489879579 CRC64;

Query Match 78.0%; Score 39; DB 1; Length 197;  
Best Local Similarity 66.7%; Pred. No. 0.5;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CXIYXPVC 11



|||||  
5 CSLIYAPVC 13

JLT 11

4 PIG

-ISK4\_PIG STANDARD; PRT; 86 AA.

P37109;

01-OCT-1994 (Rel. 30, Created)

01-OCT-1994 (Rel. 30, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Serine protease inhibitor Kazal-type 4 precursor (Peptide PEC-60).

SPINK4.

Sus scrofa (Pig).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

NCBI\_TaxID=9823;

[1]

SEQUENCE FROM N.A.

MEDLINE=93015834; PubMed=1400298;

Metsis M., Cintra A., Solfrini V., Ernfor P., Bortolotti F., Mutt V.,

Morrassutti D.G., Oestenson C.-G., Efendic S., Agerberth B., Mutt V.,

Persson H., Fuxe K.;

"Molecular cloning of PEC-60 and expression of its mRNA and peptide

in the gastrointestinal tract and immune system.";

J. Biol. Chem. 267:19829-19832(1992).

[2]

SEQUENCE OF 27-86.

TISSUE=Intestine;

MEDLINE=90046843; PubMed=2573065;

Agerberth B., Soederling-Barros J., Joernvall H., Chen Z.,

Oestenson C.G., Efendic S., Mutt V.;

"Isolation and characterization of a 60-residue intestinal peptide

structurally related to the pancreatic secretory type of trypsin

inhibitor: influence on insulin secretion.";

Proc. Natl. Acad. Sci. U.S.A. 86:8590-8594(1989).

[3]

STRUCTURE BY NMR.

MEDLINE=94254085; PubMed=8196042;

Liepinsh E., Berndt K.D., Sillard R., Mutt V., Otting G.;

"Solution structure and dynamics of PEC-60, a protein of the Kazal

type inhibitor family, determined by nuclear magnetic resonance

spectroscopy.";

J. Mol. Biol. 239:137-153(1994).

-!- FUNCTION: Inhibits the glucose-induced insulin secretion from

perfused pancreas; also plays a role in the immune system. Does

not inhibit trypsin.

-!- SUBCELLULAR LOCATION: Secreted.

-!- TISSUE SPECIFICITY: SYNTHESIZED IN DUODENAL GOBLET CELLS AND

IN MONOCYTES IN BONE MARROW AND BLOOD.

-!- SIMILARITY: Contains 1 Kazal-like domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; S46866; AAB23691.2; -.

EMBL; X67109; CAA47482.1; -.

PIR; A44041; A34427.

PDB; 1PCE; 30-APR-94.

InterPro; IPR002350; kazal.

InterPro; IPR001239; Kazal\_inhib.

Pfam; PF00050; kazal; 1.

PRINTS; PR00290; KAZALINHBTR.

SMART; SM00280; KAZAL; 1.

PROSITE; PS00282; KAZAL; 1.

3D-structure; Signal.

SIGNAL

CHAIN 1 26

SERINE PROTEASE INHIBITOR KAZAL-TYPE 4.

27 86

FT DOMAIN 35 86 KAZAL-LIKE.

FT DISULFID 37 68

FT DISULFID 46 65

FT DISULFID 54 86

FT ACT\_SITE 48 49

FT TURN 29 30 REACTIVE BOND (BY SIMILARITY).

FT STRAND 33 33

FT TURN 41 42

FT TURN 47 48

FT STRAND 53 55

FT TURN 56 57

FT STRAND 60 61

FT HELIX 64 74

FT TURN 75 75

FT STRAND 80 83

SQ SEQUENCE 86 AA; 9635 MW; 5D513142CF3A4B4D CRC64;

Query Match 76.0%; Score 38; DB 1; Length 86;

Best Local Similarity 66.7%; Pred. No. 0.36;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11

Db 46 CSRIYDPVC 54

RESULT 12

ESM1\_DROME

ID ESM1\_DROME STANDARD; PRT; 156 AA.

AC O97176;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Enhancer of split M1 protein precursor (B(spl)m1).

GN M1 OR CG8342.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=99173785; PubMed=10072784;

RA Wurmbach E., Wech I., Preiss A.;

RT "The Enhancer of split complex of Drosophila melanogaster harbors

three classes of Notch responsive genes.";

RL Mech. Dev. 80:171-180(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT THE TIME WHEN SEPARATION OF  
CC NEURAL AND EPIDERMAL PRECURSORS CELLS OCCURS. ACCUMULATES  
CC TRANSIENTLY AT THE FUSION SITES OF ANTERIOR AND POSTERIOR MIDGUT  
CC AND VERY SPECIFICALLY TO HIGH LEVELS IN THE PROVENTRICULUS OF THE  
CC EMBRYO.  
CC -!- SIMILARITY: Contains 1 Kazal-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ010167; CAB39163.1; -;  
DR EMBL; AE003754; AAF56548.1; -;  
DR FlyBase; FBgn0002578; ml.  
DR InterPro; IPR002350; kazal.  
DR Pfam; PF00050; kazal; 1.  
DR SMART; SM00280; KAZAL; 2.  
DR PROSITE; PS00282; KAZAL; FALSE NEG.  
KW Serine protease inhibitor; Differentiation; Neurogenesis; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 156 ENHANCER OF SPLIT M1 PROTEIN.  
FT DOMAIN 29 75 KAZAL-LIKE.  
SQ SEQUENCE 156 AA; 17339 MW; 5C29F73AB89F6949 CRC64;  
  
Query Match 76.0%; Score 38; DB 1; Length 156;  
Best Local Similarity 66.7%; Pred. No. 0.64;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 CXXIYXPVC 11  
DB 106 CSMIYQPVC 114  
  
RESULT 13  
IOV7\_CHICK STANDARD; PRT; 472 AA.  
AC P10184;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ovicinhibitor precursor.  
GN OIH.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-472.  
RX MEDLINE=87194792; PubMed=3571241;  
RA Scott M.J., Huckaby C.S., Kato I., Kohr W.J., Laskowski M. Jr.,  
RA Tsai M.-J., O'Malley B.W.;

RT "Ovicinhibitor introns specify functional domains as in the related  
RT and linked ovomucoid gene.";  
RL J. Biol. Chem. 262:5899-5907(1987).  
CC -!- FUNCTION: SEEMS TO HAVE AT LEAST FIVE ACTIVE INHIBITORY DOMAINS;  
CC TWO FOR TRYPSIN, TWO FOR CHYMOTRYPSIN AND ONE FOR ELASTASE.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: EGG WHITE AND PLASMA.  
CC -!- SIMILARITY: Contains 7 Kazal-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M16141; AAA48994.1; -;  
DR EMBL; M15962; AAA48994.1; JOINED.  
DR EMBL; M16127; AAA48994.1; JOINED.  
DR EMBL; M16128; AAA48994.1; JOINED.  
DR EMBL; M16129; AAA48994.1; JOINED.  
DR EMBL; M16130; AAA48994.1; JOINED.  
DR EMBL; M16131; AAA48994.1; JOINED.  
DR EMBL; M16132; AAA48994.1; JOINED.  
DR EMBL; M16133; AAA48994.1; JOINED.  
DR EMBL; M16134; AAA48994.1; JOINED.  
DR EMBL; M16135; AAA48994.1; JOINED.  
DR EMBL; M16136; AAA48994.1; JOINED.  
DR EMBL; M16137; AAA48994.1; JOINED.  
DR EMBL; M16138; AAA48994.1; JOINED.  
DR EMBL; M16139; AAA48994.1; JOINED.  
DR EMBL; M16140; AAA48994.1; JOINED.  
DR PIR; A26730; A26730.  
DR HSSP; P05586; ZOVO.  
DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 7.  
DR PRINTS; PR00290; KAZALINHBTR.  
DR SMART; SM00280; KAZAL; 7.  
DR PROSITE; PS00282; KAZAL; 7.  
KW Serine protease inhibitor; Repeat; Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 472 OVONHIBITOR.  
FT DOMAIN 26 91 KAZAL-LIKE 1.  
FT DOMAIN 92 156 KAZAL-LIKE 2.  
FT DOMAIN 157 222 KAZAL-LIKE 3.  
FT DOMAIN 223 287 KAZAL-LIKE 4.  
FT DOMAIN 288 353 KAZAL-LIKE 5.  
FT DOMAIN 354 419 KAZAL-LIKE 6.  
FT DOMAIN 412 472 KAZAL-LIKE 7.  
FT ACT\_SITE 47 48 REACTIVE BOND 1 (POSSIBLY TRYPSIN).  
FT ACT\_SITE 112 113 REACTIVE BOND 2 (POSSIBLY TRYPSIN).  
FT ACT\_SITE 178 179 REACTIVE BOND 3 (POSSIBLY TRYPSIN).  
FT ACT\_SITE 244 245 REACTIVE BOND 4 (POSSIBLY TRYPSIN).  
FT ACT\_SITE 309 310 REACTIVE BOND 5 (POSSIBLY CHYMOTRYPSIN).  
FT ACT\_SITE 375 376 REACTIVE BOND 6 (POSSIBLY CHYMOTRYPSIN  
AND ELASTASE).  
FT ACT\_SITE 434 435 REACTIVE BOND 7 (POSSIBLY CHYMOTRYPSIN  
AND ELASTASE).  
FT DISULFID 28 67  
FT DISULFID 45 64  
FT DISULFID 53 85  
FT DISULFID 94 132  
FT DISULFID 110 129  
FT DISULFID 118 150  
FT DISULFID 159 198  
FT DISULFID 176 195  
FT DISULFID 184 216  
FT DISULFID 225 264  
FT DISULFID 242 261  
FT DISULFID 250 282  
FT DISULFID 291 329

```

DISULFID 307 326
DISULFID 315 347
DISULFID 356 395
DISULFID 373 392
DISULFID 381 413
DISULFID 422 454
DISULFID 432 451
DISULFID 440 472
CONFLICT 333 333 MISSING (IN REF. 1; AA SEQUENCE).
SEQUENCE 472 AA; 51919 MW; A30211B297B800E5 CRC64;

very Match 75.0%; Score 38; DB 1; Length 472;
st Local Similarity 66.7%; Pred. No. 1.9;
atches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

3 CXXIYXPVC 11
| | | | |
373 CTMIYDPVC 381

JLT 14
_RAT
_AGR1 RAT STANDARD; PRT; 1959 AA.
P25304; Q63034;
01-MAY-1992 (Rel. 22, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Agrin precursor.
AGRN.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE OF 1-1779 AND 1799-1959 FROM N.A.
TISSUE=Embryonic spinal cord;
MEDLINE=91222570; PubMed=1851019;
Rupp F., Payan D.G., Magill-Solc C., Cowan D.M., Scheller R.H.;
"Structure and expression of a rat agrin.";
Neuron 6:811-823(1991).
[2]
SEQUENCE OF 1777-1801 FROM N.A.
MEDLINE=92407628; PubMed=1326608;
Rupp P., Oezcelik T., Linial M., Peterson K., Francke U., Scheller R.;
"Structure and chromosomal localization of the mammalian agrin gene.";
J. Neurosci. 12:3535-3544(1992).
-!- FUNCTION: Component of the basal lamina that causes the
aggregation of acetylcholine receptors and acetylcholine-esterase
on the surface of muscle fibers of the neuromuscular junction.
-!- SUBUNIT: Binds to laminin.
-!- SUBCELLULAR LOCATION: Synaptic basal lamina at the neuromuscular
junction.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=5;
Comment=Additional isoforms seem to exist. Isoforms differ in
their acetylcholine receptor clustering activity;
Name=1;
IsoId=P25304-1; Sequence=Displayed;
Name=2;
IsoId=P25304-2; Sequence=VSP_001365;
Name=3;
IsoId=P25304-3; Sequence=VSP_001366;
Name=4;
IsoId=P25304-4; Sequence=VSP_001367;
Name=5;
IsoId=P25304-5; Sequence=VSP_001368;
-!- TISSUE SPECIFICITY: Embryonic nervous system and muscle.
-!- DEVELOPMENTAL STAGE: More abundant early in development.
-!- PTM: Contains heparan sulfate chains as well as N-linked and O-
linked oligosaccharides (By similarity).
-!- SIMILARITY: Contains 9 Kazal-like domains.
-!- SIMILARITY: Contains 2 laminin EGF-like domains.
-!- SIMILARITY: Contains 4 EGF-like domains.

```

```

CC -!- SIMILARITY: Contains 1 SEA domain.
CC -!- SIMILARITY: Contains 3 laminin G-like domains.
CC -!- CAUTION: It is uncertain whether Met-1, Met-18 or Met-24 is the
CC initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M64780; AAA40703.1; -.
CC EMBL; M64780; AAA40702.1; ALT_INIT.
CC EMBL; S44194; AAB23326.1; -.
CC PIR; JH0399; AGRT.
CC HSSP; P00740; LEDM.
CC InterPro; IPR008985; ConA_like_lec_gl.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR003645; FOLN.
CC InterPro; IPR002350; Kazal.
CC InterPro; IPR002049; Laminin_EGF.
CC InterPro; IPR001791; Laminin_G.
CC InterPro; IPR000082; SEA_domain.
CC Pfam; PF00008; EGF; 4.
CC Pfam; PF00050; kazal; 9.
CC Pfam; PF00053; laminin_EGF; 2.
CC Pfam; PF00054; laminin_G; 3.
CC Pfam; PF01390; SEA; 1.
CC PRINTS; PR00011; EGFLAMININ.
CC SMART; SM00180; EGF_Lam; 2.
CC SMART; SM00274; FOLN; 8.
CC SMART; SM00280; KAZAL; 9.
CC SMART; SM00282; LamG; 3.
CC SMART; SM00200; SEA; 1.
CC PROSITE; PS00022; EGF_1; 6.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS50026; EGF_3; 4.
CC PROSITE; PS50025; LAM_G_DOMAIN; 3.
CC PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
CC PROSITE; PS50024; SEA; 1.
KW Glycoprotein; EGF-like domain; Repeat; Alternative splicing; Signal;
KW Laminin EGF-like domain; Proteoglycan; Heparan sulfate.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 1959 AGRIN.
FT DOMAIN 65 137 KAZAL-LIKE 1.
FT DOMAIN 141 212 KAZAL-LIKE 2.
FT DOMAIN 213 284 KAZAL-LIKE 3.
FT DOMAIN 287 356 KAZAL-LIKE 4.
FT DOMAIN 361 429 KAZAL-LIKE 5.
FT DOMAIN 430 494 KAZAL-LIKE 6.
FT DOMAIN 495 559 KAZAL-LIKE 7.
FT DOMAIN 563 645 KAZAL-LIKE 8.
FT DOMAIN 688 741 LAMININ_EGF-LIKE 1.
FT DOMAIN 742 788 LAMININ_EGF-LIKE 2.
FT DOMAIN 794 864 KAZAL-LIKE 9.
FT DOMAIN 1023 1145 SEA.
FT DOMAIN 1220 1258 EGF-LIKE 1.
FT DOMAIN 1263 1439 LAMININ_G-LIKE 1.
FT DOMAIN 1440 1477 EGF-LIKE 2.
FT DOMAIN 1479 1516 EGF-LIKE 3.
FT DOMAIN 1526 1708 LAMININ_G-LIKE 2.
FT DOMAIN 1709 1748 EGF-LIKE 4.
FT DOMAIN 1784 1956 LAMININ_G-LIKE 3.
FT DOMAIN 869 992 SER/THR-RICH.
FT DOMAIN 1147 1215 SER/THR-RICH.
FT DISULFID 97 116 POTENTIAL.
FT DISULFID 105 137 POTENTIAL.
FT DISULFID 171 191 POTENTIAL.
FT DISULFID 180 212 POTENTIAL.
FT DISULFID 244 263 POTENTIAL.
FT DISULFID 252 284 POTENTIAL.

```



FT DISULFID 316 335 POTENTIAL.  
FT DISULFID 324 356 POTENTIAL.  
FT DISULFID 389 408 POTENTIAL.  
FT DISULFID 397 429 POTENTIAL.  
FT DISULFID 454 473 POTENTIAL.  
FT DISULFID 462 494 POTENTIAL.  
FT DISULFID 518 538 POTENTIAL.  
FT DISULFID 527 559 POTENTIAL.  
FT DISULFID 604 624 POTENTIAL.  
FT DISULFID 613 645 POTENTIAL.  
FT DISULFID 688 700 BY SIMILARITY.  
FT DISULFID 690 707 BY SIMILARITY.  
FT DISULFID 709 718 BY SIMILARITY.  
FT DISULFID 721 739 BY SIMILARITY.  
FT DISULFID 742 754 BY SIMILARITY.  
FT DISULFID 744 761 BY SIMILARITY.  
FT DISULFID 763 772 BY SIMILARITY.  
FT DISULFID 775 786 BY SIMILARITY.  
FT DISULFID 823 843 POTENTIAL.  
FT DISULFID 832 864 POTENTIAL.  
FT DISULFID 1224 1235 BY SIMILARITY.  
FT DISULFID 1229 1246 BY SIMILARITY.  
FT DISULFID 1248 1257 BY SIMILARITY.  
FT DISULFID 1444 1455 POTENTIAL.  
FT DISULFID 1449 1465 POTENTIAL.  
FT DISULFID 1467 1476 POTENTIAL.  
FT DISULFID 1483 1494 BY SIMILARITY.  
FT DISULFID 1488 1504 BY SIMILARITY.  
FT DISULFID 1506 1515 BY SIMILARITY.  
FT DISULFID 1713 1727 BY SIMILARITY.  
FT DISULFID 1721 1736 BY SIMILARITY.  
FT DISULFID 1738 1747 BY SIMILARITY.  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 827 827 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 957 957 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 1144 1152 Missing (in isoform 2).  
FT VARSPLIC 1780 1798 Missing (in isoform 3).  
FT VARSPLIC 1788 1798 Missing (in isoform 4).  
FT VARSPLIC 1780 1797 Missing (in isoform 5).  
FT VARIANT 314 314 V -> VTCD (IN A VARIANT).  
SQ SEQUENCE 1959 AA; 208645 MW; 7FEFDFDAFF89CC31 CRC64;

Query Match 76.0%; Score 38; DB 1; Length 1959;  
Best Local Similarity 66.7%; Pred. No. 7.4;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
| | | | |  
Db 389 CSGIYDPVC 397

RESULT 15  
IOVO FRACO STANDARD; PRT; 53 AA.  
AC P05595;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ovomuroid (Fragment).  
OS Francolinus coqui (Coqui francolin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Francolinus.  
OX NCBI\_TaxID=9021;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=87157615; PubMed=3828298;  
RA Laskowski M. Jr., Kato I., Ardelt W., Cook J., Denton A., Empie M.W.,

RA Kohr W.J., Park S.J., Parks K., Schatzley B.L., Schoenberger O.L.,  
RA Tashiro M., Vichot G., Whatley H.E., Wiczorek A., Wiczorek M.;  
RT "Ovomuroid third domains from 100 avian species: isolation,  
RT sequences, and hypervariability of enzyme-inhibitor contact  
RT residues";  
RL Biochemistry 26:202-221(1987).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: AVIAN OVOMUCOID CONSIST OF THREE HOMOLOGOUS, TANDEM KAZAL  
CC FAMILY INHIBITORY DOMAINS.  
CC -!- SIMILARITY: Contains 3 Kazal-like domains.  
DR PIR; G31438; G31438.  
DR HSSP; P05586; 4OVO.  
DR InterPro; IPR002350; kazal.  
DR InterPro; IPR001239; Kazal\_inhib.  
DR Pfam; PF00050; kazal; 1.  
DR PRINTS; PR00290; KAZALINHETR.  
DR SMART; SMO0280; KAZAL; 1.  
DR PROSITE; PS00282; KAZAL; 1.  
KW Serine protease inhibitor; Glycoprotein; Repeat; Egg white.  
FT NON\_TER 1 1  
FT DOMAIN 3 53 KAZAL-LIKE 3.  
FT ACT\_SITE 15 16 REACTIVE BOND 3.  
FT DISULFID 5 35  
FT DISULFID 13 32  
FT DISULFID 21 53  
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .).  
FT NON\_TER 53 53  
SQ SEQUENCE 53 AA; 5739 MW; SEC6EBC58EDE7846 CRC64;

Query Match 74.0%; Score 37; DB 1; Length 53;  
Best Local Similarity 60.0%; Pred. No. 0.36;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11  
| | | | |  
Db 12 GCTMEYRPVC 21

Search completed: August 11, 2004, 13:08:06  
Job time : 8.38095 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 11, 2004, 13:04:18 ; Search time 35.0952 Seconds  
(without alignments)  
98.894 Million cell updates/sec

le: US-10-661-207-2  
fect score: 50  
ence: 1 XGCXXIYXPVC 11

ring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

ched: 1017041 seqs, 315518202 residues

al number of hits satisfying chosen parameters: 1017041

imum DB seq length: 0  
imum DB seq length: 2000000000

:-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

ibase :

- SPTREMBL 25:\*
- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

alt	Score	Query Match	Length	DB	ID	Description
1	43	86.0	155	5	Q8T926	Q8t926 tetrahymena
2	38	76.0	730	5	Q9VNL6	Q9vnl6 drosophila
3	37	74.0	56	6	Q29185	Q29185 sus scrofa
4	37	74.0	2026	4	O00468	O00468 homo sapien
5	36	72.0	94	5	Q9VHK7	Q9vhk7 drosophila
6	36	72.0	145	5	Q9VE57	Q9ve57 drosophila
7	36	72.0	1699	5	O02037	O02037 bdelloura c
8	35	70.0	140	12	O36986	O36986 lily sympto
9	35	70.0	140	12	Q83090	Q83090 lily sympto
10	35	70.0	222	5	Q95P16	Q95p16 triatoma in
11	35	70.0	385	4	Q92659	Q92659 homo sapien
12	35	70.0	406	16	Q8F6E2	Q8f6e2 leptospira
13	35	70.0	724	10	Q9LKT8	Q9lkt8 arabidopsis
14	34	68.0	82	12	Q91GK0	Q91gk0 epiphyas po
15	34	68.0	82	12	Q7TLU0	Q7tlu0 choristoneu
16	34	68.0	93	12	O73509	O73509 potato viru

17	34	68.0	93	12	O73529	O73529 potato viru
18	34	68.0	93	12	O73531	O73531 potato viru
19	34	68.0	93	12	O73512	O73512 potato viru
20	34	68.0	93	12	O73508	O73508 potato viru
21	34	68.0	93	12	O93139	O93139 potato viru
22	34	68.0	93	12	O73514	O73514 potato viru
23	34	68.0	93	12	O73527	O73527 potato viru
24	34	68.0	94	12	O41486	O41486 potato viru
25	34	68.0	94	12	O41484	O41484 potato viru
26	34	68.0	94	12	O73525	O73525 potato viru
27	34	68.0	94	12	O86541	O86541 potato viru
28	34	68.0	237	5	Q25196	Q25196 hydra atten
29	34	68.0	321	10	Q84JM1	Q84jml arabidopsis
30	34	68.0	335	10	Q9C559	Q9c559 arabidopsis
31	34	68.0	341	5	Q24329	Q24329 drosophila
32	34	68.0	341	5	Q9VU62	Q9vu62 drosophila
33	34	68.0	366	5	Q8SZW6	Q8szw6 drosophila
34	34	68.0	405	11	Q8CB11	Q8c811 mus musculu
35	34	68.0	746	10	Q9SJJ15	Q9sj15 arabidopsis
36	34	68.0	1120	4	Q9BX34	Q9bx34 homo sapien
37	34	68.0	1164	4	Q86XP2	Q86xp2 homo sapien
38	34	68.0	1220	4	Q86XP1	Q86xp1 homo sapien
39	33	66.0	169	5	Q9U5B4	Q9usb4 caenorhabdi
40	33	66.0	274	16	Q7UI56	Q7ui56 rhodopirell
41	33	66.0	279	5	Q8I600	Q8i600 plasmodium
42	33	66.0	370	16	Q8A715	Q8a715 bacteroides
43	33	66.0	757	10	Q9SPJ0	Q9spj0 oryza sativ
44	33	66.0	759	10	Q941S0	Q94ls0 avena strig
45	33	66.0	765	10	Q9SLP9	Q9slp9 luffa cylin

ALIGNMENTS

RESULT 1

Q8T926 ID Q8T926 PRELIMINARY; PRT; 155 AA.  
AC Q8T926;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Kazal-type proteinase inhibitor 1.  
OS Tetrahymena thermophila.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;  
OC Tetrahymenina; Tetrahymena.  
OX NCBI\_TaxID=5911;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Turkewitz A.P., Haddad A.;  
RT "A non-architectural cargo protein in Tetrahymena thermophila dense  
RT core secretory granules."  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY075151; AAL79514.1; --  
DR InterPro; IPR002350; Kazal.  
DR SMART; SM00280; KAZAL; 1.  
SQ SEQUENCE 155 AA; 17608 MW; E438CC6E2110786B CRC64;

Query Match 86.0%; Score 43; DB 5; Length 155;  
Best Local Similarity 70.0%; Pred. No. 0.29;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCXXIYXPVC 11  
|||  
Db 31 GCPEIYSPVC 40

RESULT 2

Q9VNL6 ID Q9VNL6 PRELIMINARY; PRT; 730 AA.  
AC Q9VNL6;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE CG1077 protein.  
 GN CG1077.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Hernandez J.R., Houck J.,  
 RA Harris N.L., Harvey D., Heiman T.J., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Hostin D., Houston K.A., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Jalali M., Kalush F., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Liang Y., Lin X.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu L.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AF003600; AAF51914.1; -.  
 DR FlyBase; FBgn0037405; CG1077.  
 DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
 DR InterPro; IPR002350; kazal.  
 DR Pfam; PF00050; kazal; 3.  
 DR SMART; SM00280; KAZAL; 3.  
 KW Protease inhibitor; Serine protease inhibitor.  
 SQ SEQUENCE 730 AA; 80473 MW; 3290B882D74E5A09 CRC64;

Query Match 76.0%; Score 38; DB 5; Length 730;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
 |||||  
 Db 150 CTRIYRPVC 158

## RESULT 3

Q29185  
 ID Q29185 PRELIMINARY; PRT; 56 AA.  
 AC Q29185;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Pancreatic secretory trypsin inhibitor (Fragment).  
 OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Small intestine;  
 RX MEDLINE=96327607; PubMed=8672129;  
 RA Winteroe A.K., Fredholm M., Davies W.;  
 RT "Evaluation and characterization of a porcine small intestine cDNA  
 RT library."  
 RL Mamm. Genome 7:509-517(1996).  
 DR EMBL; F14496; CAA23088.1; -.  
 DR HSSP; P37109; LPCE.  
 DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
 DR InterPro; IPR002350; kazal.  
 DR InterPro; IPR001239; Kazal\_inhib.  
 DR Pfam; PF00050; kazal; 1.  
 DR PRINTS; PR00290; KAZALINHBTR.  
 DR SMART; SM00280; KAZAL; 1.  
 DR PROSITE; PS00282; KAZAL; 1.  
 FT NON\_TER 1  
 FT NON\_TER 56  
 SQ SEQUENCE 56 AA; 6041 MW; 39ABD7169DF16D25 CRC64;

Query Match 74.0%; Score 37; DB 6; Length 56;  
 Best Local Similarity 60.0%; Pred. No. 2;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11  
 |||||  
 Db 15 GCPKIYNPXC 24

## RESULT 4

Q00468  
 ID Q00468 PRELIMINARY; PRT; 2026 AA.  
 AC Q00468;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE AGRIN precursor (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Groffen A.J.A., Buskens C.A.F., Van Kuppevelt T.H.M.S.M.,

RA Veerkamp J.H., Monnens L.A.H., van den Heuvel L.P.W.J.;

RL Eur. J. Biochem. 0:0-0(1998).

RN [2]

RP SEQUENCE OF 1-153 FROM N.A.

RX MEDLINE=96224170; PubMed=8617505;

RA Lennon G., Auffray C., Polymeropoulos M., Soares M.B.;

RT "The I.M.A.G.E. Consortium: an integrated molecular analysis of

RT genomes and their expression."

RL Genomics 33:151-152(1996).

RN [3]

RP SEQUENCE OF 1-153 FROM N.A.

RA Denzer A.J., Brandenberger R., Gesemann M., Chiquet M., Ruegg M.A.;

RL J. Cell Biol. 0:0-0(0).

DR EMBL; AF016903; AAC39776.1; -.

DR EMBL; U84406; AAB52917.1; -.

DR HSSP; P00740; IEDM.

DR Genew; HGNC:329; AGRN.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR004850; Agrin\_NtA.

DR InterPro; IPR006209; EGF\_Like.

DR InterPro; IPR003645; FOLN.

DR InterPro; IPR002350; kazal.

InterPro; IPR002049; Laminin\_EGF.  
InterPro; IPR001791; Laminin\_G.  
InterPro; IPR000082; SEA domain.  
InterPro; IPR008993; TIMP like.  
InterPro; IPR001455; UPF0033.  
Pfam; PF00008; EGF; 4.  
Pfam; PF00050; kazal; 9.  
Pfam; PF00053; laminin\_EGF; 2.  
Pfam; PF00054; laminin\_G; 3.  
Pfam; PF03146; NtA; 1.  
Pfam; PF01390; SEA; 1.  
PRINTS; PR00011; EGFLAMININ.  
SMART; SM00180; EGF\_Lam; 2.  
SMART; SM00274; FOLN; 5.  
SMART; SM00280; KAZAL; 9.  
SMART; SM00282; LamG; 3.  
SMART; SM00200; SEA; 1.  
PROSITE; PS00022; EGF\_1; 6.  
PROSITE; PS01186; EGF\_2; 1.  
PROSITE; PS01248; LAMININ\_TYPE\_EGF; 1.  
PROSITE; PS50025; LAM\_G\_DOMAIN; 3.  
PROSITE; PS50024; SEA; 1.  
PROSITE; PS01148; UPF0033; 1.  
EGF-like domain; Laminin EGF-like domain; Signal.  
NON\_TER 1 1  
SIGNAL <1 10 POTENTIAL.  
CHAIN 11 2026 AGRIN.  
SEQUENCE 2026 AA; 212881 MW; 4AB0EE710CD4B8EF CRC64;  
  
Query Match 74.0%; Score 37; DB 4; Length 2026;  
Best Local Similarity 55.6%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
3 CXXIYXPVC 11  
| : | | |  
475 CSSLYDPVC 483

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003680; AAF54297.1; -.  
DR EMBL; AY118777; AAM50637.1; -.  
DR FlyBase; FBgn0040532; CG8369.  
SQ SEQUENCE 94 AA; 10163 MW; 1A8EA3359E520788 CRC64;  
  
Query Match 72.0%; Score 36; DB 5; Length 94;  
Best Local Similarity 55.6%; Pred. No. 52;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 CXXIYXPVC 11  
| : | | |  
Db 37 CGEVEPVC 45  
  
RESULT 6  
Q9VE57 PRELIMINARY; PRT; 145 AA.  
ID Q9VE57 PRELIMINARY; PRT; 145 AA.  
AC Q9VE57;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CG7695 protein.  
GN CG7695.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,



RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,  
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003723; AAF55572.1; -.  
DR FlyBase; FBgn0038631; CG7695.  
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
DR InterPro; IPR002350; kazal.  
DR Pfam; PF00050; kazal; 1.  
DR SMART; SM00280; KAZAL; 1.  
KW Protease inhibitor; Serine protease inhibitor.  
SQ SEQUENCE 145 AA; 16581 MW; 4A52F2795CC67288 CRC64;

Query Match 72.0%; Score 36; DB 5; Length 145;  
Best Local Similarity 66.7%; Pred. No. 7.8;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CXXIYXPVC 11  
| | | | |  
Db 95 CPDIYDPVC 103

## RESULT 7

ID 002037 PRELIMINARY; PRT; 1699 AA.

AC 002037;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Voltage-gated sodium channel homolog Bdnal.  
GN BDNAL.

OS Bdeloura candida.  
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;  
OC Maricola; Bdellouroidea; Bdellouridae; Bdelloura.  
OX NCBI\_TaxID=46766;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97445581; PubMed=9300466;  
RA Jeziorski M.C., Greenberg R.M., Anderson P.A.V.;  
RT "Cloning of a putative voltage-gated sodium channel from the  
RT turbellarian flatworm Bdelloura candida.";  
RL Parasitology 115:289-296(1997).

DR EMBL; U93074; AAC63049.1; -.

DR PIR; T31340; T31340.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0001518; C:voltage-gated sodium channel complex; IEA.

DR GO; GO:0005261; F:cation channel activity; IEA.

DR GO; GO:0005248; F:voltage-gated sodium channel activity; IEA.

DR GO; GO:0006812; P:cation transport; IEA.

DR GO; GO:0006814; P:sodium ion transport; IEA.  
DR InterPro; IPR001682; Ca/Na\_pore.  
DR InterPro; IPR002111; Cat\_channel\_TrpL.  
DR InterPro; IPR001064; Crystallin.  
DR InterPro; IPR005821; Ion\_trans.  
DR InterPro; IPR005820; M+channel\_nlg.  
DR InterPro; IPR001696; Na\_channel.  
DR Pfam; PF00520; ion\_trans; 4.  
DR PRINTS; PR00170; NACHANNEL.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
KW Ionic channel; Transmembrane.  
SQ SEQUENCE 1699 AA; 197754 MW; F985DSF34D9CE62A CRC64;

Query Match 72.0%; Score 36; DB 5; Length 1699;

Best Local Similarity 55.6%; Pred. No. 73;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CXXIYXPVC 11  
| | | | |  
Db 469 CDDIYQPIC 477

## RESULT 8

O36986

ID O36986 PRELIMINARY; PRT; 140 AA.

AC O36986;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE 16kDa unknown protein.

OS Lily symptomless virus (LSV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlaviruses.

OX NCBI\_TaxID=12173;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LSV-Ko;

RA Ryu K.H., Ahn H.I., Kim S.J., Park W.M.;

RT "The nucleotide sequence of 3'-terminal region of lily symptomless

RT virus (LSV) Korean strain (LSV-Ko).";

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF015286; AAB70213.1; -.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR002568; Carla\_C4.

DR Pfam; PF01623; Carla\_C4; 1.

SQ SEQUENCE 140 AA; 16113 MW; 5FA2069B9EDDA927 CRC64;

## Query Match

70.0%; Score 35; DB 12; Length 140;

Best Local Similarity 55.6%; Pred. No. 12;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CXXIYXPVC 11  
| | | | |  
Db 86 CYRVYPPVC 94

## RESULT 9

Q83090

ID Q83090 PRELIMINARY; PRT; 140 AA.

AC Q83090;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE 16 kDa protein.

OS Lily symptomless virus (LSV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlaviruses.

OX NCBI\_TaxID=12173;

RN [1]

RP SEQUENCE FROM N.A.

RA Kim J.K., Kim Y.H., Lee S.Y.;

RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U43905; AAA85904.1; -.

DR GO; GO:0003676; F:nucleic acid binding; IEA.



GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 InterPro; IPR002568; Carla\_C4.  
 Pfam; PF01623; Carla\_C4; 1.  
 SEQUENCE 140 AA; 16051 MW; A190FA0BDE82A639 CRC64;  
 Query Match 70.0%; Score 35; DB 12; Length 140;  
 Best Local Similarity 55.6%; Pred. No. 12;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 3 CXXIYXPVC 11  
 | : |||  
 86 CYRVYPPVC 94  
 ULT 10  
 P16  
 Q95P16 PRELIMINARY; PRT; 222 AA.  
 Q95P16;  
 01-DEC-2001 (TrEMBLrel. 19, Created)  
 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 Thrombin inhibitor infestin (Fragment).  
 Triatoma infestans (Assassin bug).  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;  
 Panheteroptera; Cimicomorpha; Reduviidae; Triatominae; Triatoma.  
 NCBI\_TaxID=30076;  
 [1]  
 SEQUENCE FROM N.A.  
 TISSUE=anterior midgut;  
 Campos I.T.N., Amino R., Sampaio C.A.M., Auerswald E.A., Friedrich T.,  
 Lemaire H.-G., Schenkman S., Tanaka A.S.;  
 "Infestin, a novel thrombin inhibitor present in Triatoma infestans  
 midgut, a Chagas disease vector: cloning, expression and  
 characterization.";  
 Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; AF360846; AAK57342.1; -.  
 HSSP; P01001; 1BUS.  
 GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
 InterPro; IPR002350; kazal.  
 InterPro; IPR001239; Kazal\_inhib.  
 Pfam; PF00050; kazal; 4.  
 PRINTS; PR00290; KAZALINHETR.  
 SMART; SM00280; KAZAL; 4.  
 PROSITE; PS00282; KAZAL; 4.  
 NON\_TER 1  
 CHAIN 1 >104 THROMBIN INHIBITOR INFESTIN.  
 NON\_TER 222  
 SEQUENCE 222 AA; 24370 MW; C9185168B4D41AC3 CRC64;  
 Query Match 70.0%; Score 35; DB 5; Length 222;  
 Best Local Similarity 55.6%; Pred. No. 18;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 3 CXXIYXPVC 11  
 | : |||  
 117 CTMYKPVK 125  
 ULT 11  
 659  
 Q92659 PRELIMINARY; PRT; 385 AA.  
 Q92659; Q9ULB0;  
 01-FEB-1997 (TrEMBLrel. 02, Created)  
 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 Haptoglobin-related protein.  
 HPR OR A-259H10.2.  
 Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=97101114; PubMed=8945641;  
 RA Tabak S., Lev A., Valansi C., Shalitin C.;  
 RT "Transcriptionally active haptoglobin-related (Hpr) gene in Hepatoma  
 G2 and leukemia molt-4 cells.";  
 RL DNA Cell Biol. 15:1001-1007(1996).  
 RN [2]  
 RP SEQUENCE OF 40-385 FROM N.A.  
 RX MEDLINE=99425270; PubMed=10493829;  
 RA Loftus B.J., Kim U.J., Sneddon V.P., Kalush F., Brandon R.,  
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,  
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,  
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
 RT "Genome duplications and other features in 12 Mb of DNA sequence from  
 human chromosome 16p and 16q.";  
 RL Genomics 60:295-308(1999).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 DR EMBL; X89214; CAA61501.1; -.  
 DR EMBL; AC004682; AAC27433.1; -.  
 DR HSSP; P08709; 1FAK.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR008292; Pept\_S1\_HapGnph.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF00084; sushi; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PIRSF; PIRSF001137; Haptoglobin; 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT VARIANT 376 376 H -> D.  
 SQ SEQUENCE 385 AA; 43077 MW; 8445DD27A496188B CRC64;  
 Query Match 70.0%; Score 35; DB 4; Length 385;  
 Best Local Similarity 55.6%; Pred. No. 30;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 CXXIYXPVC 11  
 | : |||  
 DB 8 CVCVMPVC 16  
 RESULT 12  
 Q8F6E2 PRELIMINARY; PRT; 406 AA.  
 AC Q8F6E2;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN LA1366.  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 OX NCBI\_TaxID=173;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE011316; AAN48565.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 406 AA; 47362 MW; D088CDD184E7E93E CRC64;  
 Query Match 70.0%; Score 35; DB 16; Length 406;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11  
|| : ||  
Db 247 GCELYIPSC 256

## RESULT 13

Q9LKT8 PRELIMINARY; PRT; 724 AA.  
AC Q9LKT8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN T32B20.I.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA WashU;  
RT "The A. thaliana Genome Sequencing Project.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBSJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Wilson R.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AF262041; AAF67366.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 724 AA; 83057 MW; D75EC74414C20317 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 724;  
Best Local Similarity 40.0%; Pred. No. 54;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GCXXIYXPVC 11  
|| : ||  
Db 698 GCNAVHIPC 707

## RESULT 14.

Q91GK0 PRELIMINARY; PRT; 82 AA.  
ID Q91GK0;  
AC Q91GK0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Epiphyas postvittana nucleopolyhedrovirus.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=70600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hyink O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., Drake K.,  
RA Ward V.K.;  
RT "The complete sequence of the Epiphyas postvittana  
nucleopolyhedrovirus genome.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AY043265; AAK85615.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 82 AA; 9395 MW; 32D80D96EC0324AB CRC64;

Query Match 68.0%; Score 34; DB 12; Length 82;  
Best Local Similarity 44.4%; Pred. No. 12;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 CXXIYXPVC 11  
|| : ||  
Db 74 CSALYYPLC 82

## RESULT 15

Q7TLU0 PRELIMINARY; PRT; 82 AA.  
ID Q7TLU0;  
AC Q7TLU0;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Choristoneura fumiferana nuclear polyhedrosis virus (CfMNPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=208973;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93033705; PubMed=1413988;  
RA Lee H.Y., Arif B., Dobos P., Krell P.;  
RT "Identification of bent DNA and ARS fragments in the genome of  
RT Choristoneura fumiferana nuclear polyhedrosis virus.";  
RL Virus Res. 24:249-264(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95297142; PubMed=7778276;  
RA Xie W.D., Arif B., Dobos P., Krell P.J.;  
RT "Identification and analysis of a putative origin of DNA replication  
RT in the Choristoneura fumiferana multinucleocapsid nuclear polyhedrosis  
RT virus genome.";  
RL Virology 209:409-419(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95297155; PubMed=7778286;  
RA Liu J.J., Carstens E.B.;  
RT "Identification, localization, transcription, and sequence analysis of  
RT the Choristoneura fumiferana nuclear polyhedrosis virus DNA polymerase  
RT gene.";  
RL Virology 209:538-549(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96030854; PubMed=7595348;  
RA Barrett J.W., Krell P.J., Arif B.M.;  
RT "Characterization, sequencing and phylogeny of the ecdysteroid UDP-  
RT glucosyltransferase gene from two distinct nuclear polyhedrosis  
RT viruses isolated from Choristoneura fumiferana.";  
RL J. Gen. Virol. 76:2447-2456(1995).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96183379; PubMed=8610448;  
RA Qiu W., Liu J.J., Carstens E.B.;  
RT "Studies of Choristoneura fumiferana nuclear polyhedrosis virus gene  
RT expression in insect cells.";  
RL Virology 217:564-572(1996).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400202; PubMed=8806578;  
RA Liu J.J., Carstens E.B.;  
RT "Identification, molecular cloning, and transcription analysis of the  
RT Choristoneura fumiferana nuclear polyhedrosis virus spindle-like  
RT protein gene.";  
RL Virology 223:396-400(1996).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20276145; PubMed=10814576;  
RA Lapointe R., Back D.W., Ding Q., Carstens E.B.;  
RT "Identification and molecular characterization of the Choristoneura  
RT fumiferana multicapsid nucleopolyhedrovirus genomic region encoding  
RT the regulatory genes pkip, p47, lef-12, and gta.";  
RL Virology 271:109-121(2000).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21854555; PubMed=11864738;  
RA Carstens E.B., Liu J.J., Dominy C.;  
RT "Identification and molecular characterization of the baculovirus

CfMNPV early genes: ie-1, ie-2 and pe38.";  
Virus Res. 83:13-30(2002).  
[9]  
SEQUENCE FROM N.A.  
de Jong J.G., Dominy C.N., Lauzon H.A., Arif B.M., Carstens E.B.,  
Krell P.J.;  
"Complete Genome of Choristoneura fumiferana Multiple  
Nucleopolyhedrovirus.";  
Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; AF512031; AAP29838.1; --  
Hypothetical protein.  
SEQUENCE 82 AA; 9594 MW; BEBDEEAD779830EF CRC64;

Very Match 69.0%; Score 34; DB 12; Length 82;  
1st Local Similarity 44.4%; Pred. No. 12;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

3 CXXIYXPVC 11  
| : | | : |  
74 CSALYYPLC 82

Job completed: August 11, 2004, 13:09:25  
Time : 36.0952 secs

**ST AVAILABLE COPY**